Stretch activation in muscle: a Ca^{2+} independent mechanism?

Thesis

UNRESTRICTED

Stretch activation in muscle: a Ca^{2+} independent mechanism?

Gian Felice De Nicola

A thesis submitted to The Open University as partial fulfilment of the requirements for the Degree of Doctor of Philosophy

November 2007

Division of Protein Structure
National Institute for Medical Research
The Ridgeway, Mill Hill
London NW7 1AA

DATE OF SUBMISSION 01 DECEMBER 2007
DATE OF AWARD 20 MARCH 2008

ABSTRACT

The aim of this thesis has been the understanding at the molecular level of stretch activation in muscles. Stretch activation is a feature common to all types of muscle. Such a mechanism is particularly important in cardiac muscles in vertebrates and in indirect flight muscles in insects. Here, the three-dimensional structure of *Lethocerus* troponin C isoform (F1), recently discovered to be crucial in regulating stretch-activated contraction, is presented. The interactions between F1 and the other subunits of the *Lethocerus* troponin trimer are also analyzed. Finally a model for stretch activation in muscle, based on the experimental data produced during this thesis, is proposed.

ACKNOWLEDGEMENT

I would like to thank Geoff Kelly for teaching me how to use the NMR spectrometer and how to interpret NMR data, Steve Martin for the fluorescence and CD experiment and for reading and correcting my thesis, Belinda Bullard for providing me with the TnH and TnT samples and the F1 clone, for answering all my questions on muscle contraction and for reading and correcting my thesis. I would like also to thank Rajesh Menon and John McCormick for teaching me the molecular biology and biochemistry that was needed to carry forward my projects and the other members of the lab for the jovial atmosphere. Last and not least I would like to thank my supervisor Annalisa Pastore for giving me the opportunity to do my PhD in her group.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

1D	mono dimensional
2D	bi dimensional
3D	three dimensional
5,5-Br_2BAPTA	5,5'dibromo-1,2-bis(2-aminophenoxy)ethane-N,N,N',N'-tetraacetic acid
Apo F1	calcium free *lethocerus* troponin C (isoform 1)
ATP	adenosine triphosphate
BMRB	biological magnetic resonance data bank
CD	circular dichroism
E.Coli	Escherichia coli
EDTA	ethylenediamminetetraacetic acid
EM	energy minimization
F1	*lethocerus* troponin C (isoform 1)
F2	*lethocerus* troponin C (isoform 2)
GST	glutathione S-transferase
His_6	hexahistidine tag
Holo F1	calcium saturated *lethocerus* troponin C (isoform 2)
HSQC	heteronuclear single quantum coherence
IFM	indirect flight muscle
IPTG	isopropyl β-D-thiogalactopyranoside
iRMSD	root mean square deviation at the interface
LB	Luria-Bertani
Ni-NTA	Ni^{2+}-nitrilotriaceticacid
NMR	nuclear magnetic resonance
NOE	nuclear Overhauser effect
NOESY	nuclear Overhauser effect spectroscopy
OD_{600}	optical density measured at 600nm
PDB	protein data bank
RMSD	root mean square deviation
SR	sarcoplasmic reticulum

TAD-SA	torsion angle space
TEV	tobacco-etch-virus
Tn	troponin
TnC	troponin C
TnH	troponin H
TnI	troponin I
TnT	troponin T
TOCSY	total correlated spectroscopy

CHAPTER 1

INTRODUCTION

1.1 Muscle and the sarcomere

The work presented in this thesis involves the study of the three-dimensional structure of the *Lethocerus* troponin C isoform (F1) recently discovered to be crucial in regulating stretch-activated contraction.

Movement is a typical feature of most living creatures. For most living species the ability to move is strictly related to the chances of survival and for most species the basic mechanism that causes movement, in whatever form it happens (i.e. running, swimming, flying) is muscle contraction.

Muscle can generally be divided into two broad categories: striated and smooth. The terms derived from the way the two types of muscle appear when analyzed through an optical microscope. Vertebral skeletal, indirect flight muscles (IFM) (the one which regulates asynchronous flight in insects) and cardiac muscle belong to the striated category. The muscles found in the bladder, abdominal cavity, respiratory and gastro-intestinal tracts and in the eye iris belong to the smooth category. Most of the work reported in this thesis refers to striated muscles. Therefore from now on, unless otherwise stated, the term muscle will be used to refer to the striated muscles.

Muscle fibres are single cells made of many myofibrils which extend the length of the cell. The myofibril is organized into regular repeated units along its length called sarcomeres. The sarcomere is the basic contractile unit in the muscle. Cardiac, skeletal and IFM muscle are defined as "striated" because they contain sarcomeres. They are packed into highly-regular arrangements giving the typical striated appearance when looked at by microscopy. In skeletal muscles, myofibrils are arranged in a more regular and parallel fashion compared to cardiac muscle where they branch and connect with each other at irregular angles. This is the reason why some textbooks do not define cardiac muscle as striated.

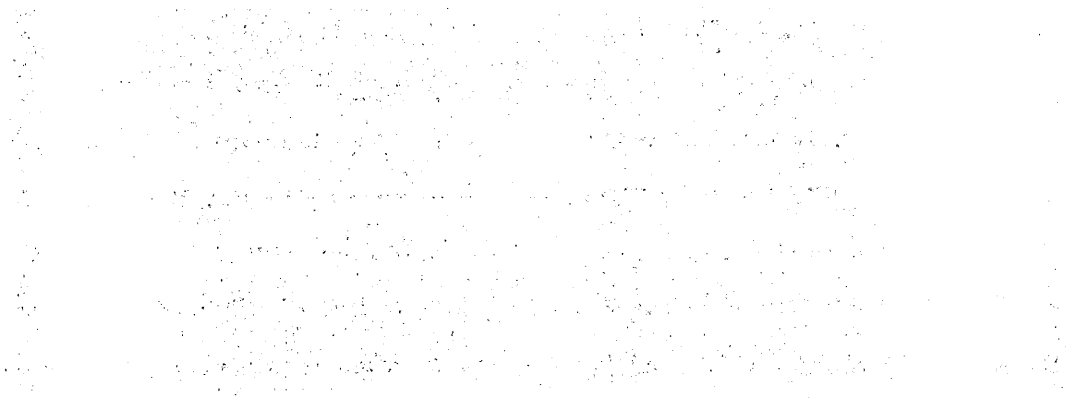
The architecture of the sarcomere is based on two types of filaments: the thick filament and the thin filament. (Squire, 1981; Cullen and Landon, 1994). The thick filament is composed mainly of myosin molecules, which are dimers with a molecular weight of 420 KDa .These molecules comprise a 1500 Å two-chain coiled-coil α-helical rod at the end of which are two similar globular 'heads' that can interact with actin. (Huxley 1969; Gergeley, 1994, Holmes, 1995) The myosin molecules are arranged so that the 'heads'

are always directed away from the mid-point of the filament. The mid-point in vertebrate sarcomere is defined by the M-line. (See fig.1.1.2) The thin filament is mainly composed of actin filaments, tropomyosin and troponin. The actin filaments are helical assemblies of actin monomers. Their filamentous structure appears as two twisted strands of actin monomers. The filament structure repeats axially after 720 . (Tobacman *et al.,* 1996) The actin monomers from adjacent sarcomeres have opposite polarity and overlap in the Z-disk where they are cross linked by alpha-actinin, the major component of the Z-disk. (See fig.1.1.2) Tropomyosin molecules are two-chain coiled-coil α- helices of length of about 400 that link end-to-end to form long strands, two of which lie along the actin helices. Each tropomyosin is associated with one globular troponin complex (molecular weight 69.5 KDa) comprising three separate chains: troponin C (molecular weight 18 KDa): the calcium binding subunit, troponin I (molecular weight 21KDa): the inhibitory subunit and troponin T (30.5KDa): the anchoring subunit.(Greaser and Gergely, 1971 Ohtsuki *et al.,* 1986, Philips and Chako, 1996)

In vertebrates an additional filament system is formed by titin which is associated with the thick filament. Vertebrate titin is a structural elastic filament (Labeit and Kolmerer, 1995). Titin is so far the largest protein known. Its molecular weight is of about 3MDa and is about 1μm long. (Trinik, 1996) It spans half of the sarcomere from the Z-disk (N-terminus) to the the M line (C-terminus). The main part of the structure consists of domains that are homologous to fibronectin type III or to immunoglobulin structure. (Improta *et al.,*1996)

This does not apply to IFM sarcomere which does not have a 3MDa titin. The elastic properties of asynchronous flight muscle are largely determined by the presence of more than one elastic protein associated with both the thick and thin filament. In IFM the function of titin is divided between two or more smaller proteins: kettin, isoform of Sls, projectin and flightin none of which extend the full length of the sarcomere. The net effect of such extra protein filaments is to reinforce both the thin and thick filament in the regions that are most exposed to strain during flight. (Bullard *et al.,* 2005)

The Z and M lines define respectively the beginning and the middle of the sarcomere. The major component of the Z-disk is α-actinin (Vigoreaux *et al.,*1994). It has a rod like appearance. The rods consist of repeats of single chain folding back onto itself to give an



anti-parallel three chain coiled-coil α-helical structure. The M-line consists of many proteins such as myomesin, M-protein and creatin kinase. (Luther and Squire, 1980)

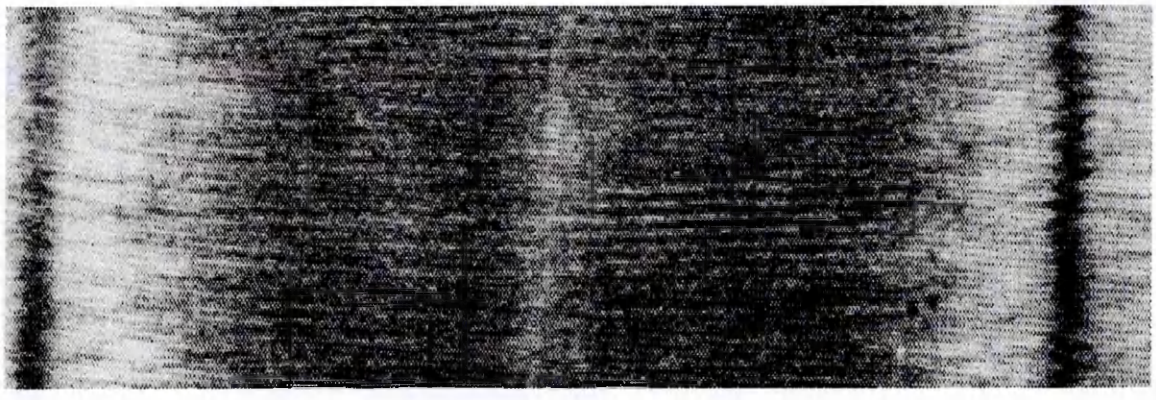


Fig.1.1.1 Electron microscopy image of the sarcomere. The dark vertical bands on the side are the Z-disk lines the paler one in the middle is the M line

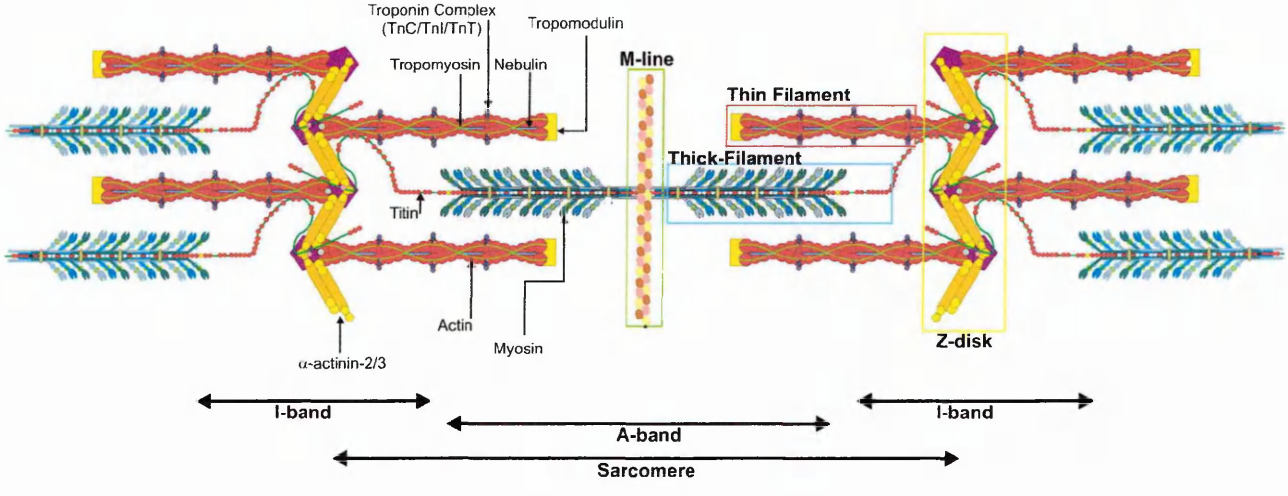


Fig.1.1.2 Schematic representation of the vertebrate skeletal sarcomere where all the major components are labelled.

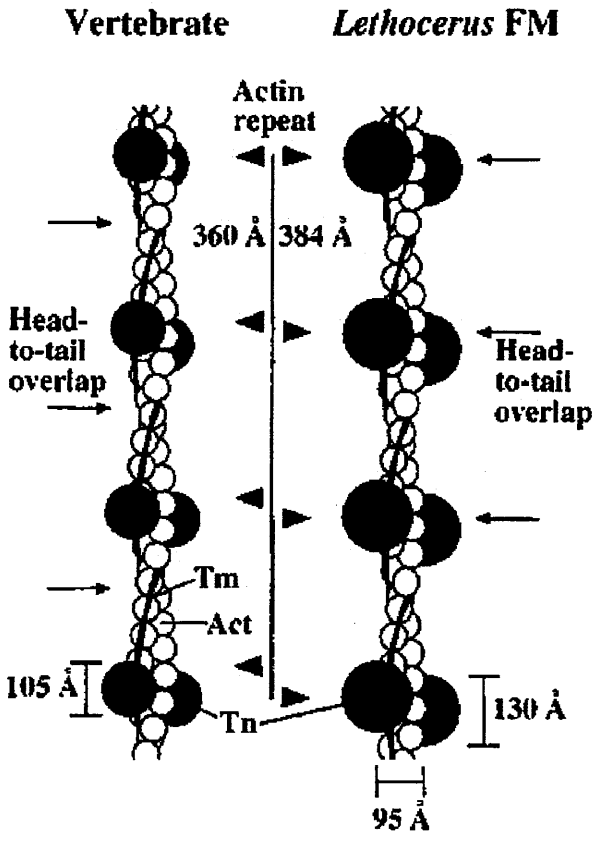


Fig.1.1.3 Differences between the vertebrate and *Lethocerus* thin filament. The size and position of the troponin trimer in the two systems is different. In *Lethocerus* the troponin trimer is located where the head to tail in the coil-coiled structure of tropomyosin overlap whereas in vertebrates the head to tail overlap does not coincide with the location of the troponin trimer. The spacing between successive trimers is also different.

Contraction, according to the sliding filament hypothesis (Huxley, 1969), is caused by the sliding of the myosin filament on the actin filament. The contraction process causes the sarcomere to shorten and produce tension. For sliding to happen, the heads of the myosin molecules have to interact with actin to form a cross bridge. Contraction starts when myosin undergoes a conformational change coupled to the hydrolysis of ATP. The conformational change causes the myosin to act as a lever arm that produces the sliding of the two filaments. In such a model actin acts as a rigid template on which the myosin can slide. In vertebrates titin limits the range of extension of the sarcomere during development of tension and contributes to the passive stiffness of the muscle keeping the M to Z distance constant. (Holmes *et al.*, 1990, Holmes, 1995)

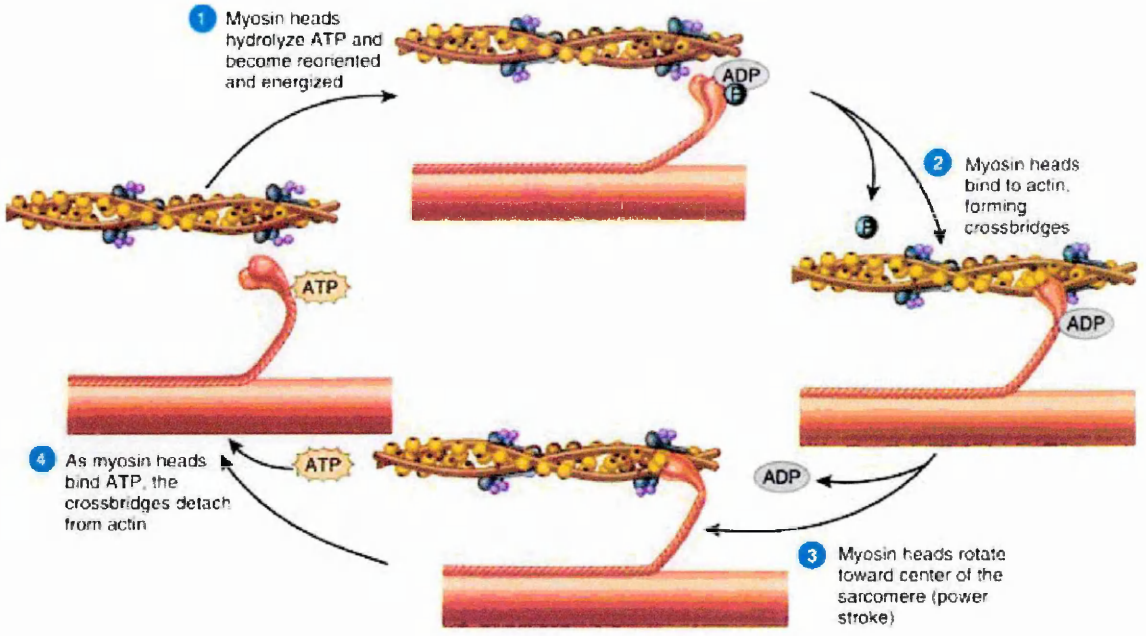


Fig.1.1.3 Schematic representation of the sliding mechanism coupled to the ATP cycle. The picture was taken from the web-site http://pharyngula.org .The muscle contraction cycle is explained. The cycle can continue as long as both ATP and Ca^{2+} are present.

1.2 Troponin isoforms in vertebrates and insect indirect flight muscles

Troponin (Tn) is a protein heterotrimer in the thin filament. The three subunits are: Troponin T (TnT), troponin C (TnC), and troponin I (TnI). TnC is the calcium sensor subunit in the trimer and troponin I is the inhibitory subunit. The role of TnT is to anchor the whole troponin trimer to the thin filament surface. The anchoring is achieved through an interaction with the tropomyosin coiled coil which itself interacts with actin. These features are common to all known troponins. (Greaser and Gergely, 1971 Gordon *et al.*, 2000)

An N terminal extension in the cardiac TnI subunit and the number of calcium binding sites in the TnC subunit (four for skeletal muscle and three for the cardiac one) are the main differences between the skeletal and cardiac isoforms in vertebrates. The skeletal and cardiac troponin trimer X ray structures are the only available 3D structure of the troponin trimer. (Takeda *et al.*,2003; Vinogradova *et al.*,2005)

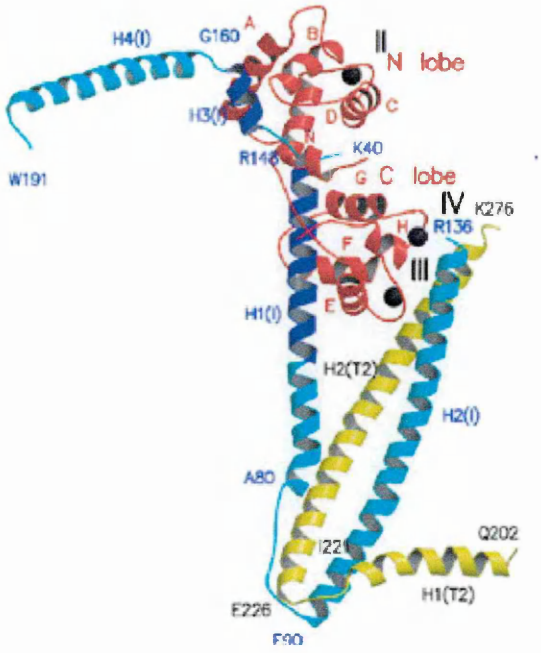


Fig.1.2.1 The X-ray structure of cardiac troponin taken from Takeda *et al.*, (2003). TnT, TnI and TnC are coloured in yellow, blue and red, respectively. The interaction between TnT and TnI through a coiled coil is visible. The roman numbering indicates the calcium binding sites on TnC

In insect IFM there are two troponin isoforms which differ in the TnC subunit. The two TnC subunits are designated F1 and F2: F1 binds just a single calcium ion whereas F2 binds to two.

IFM troponin H the *Lethocerus* ortholog of TnI also differs from the cardiac and skeletal forms in having a long alanine and proline rich C-terminal extension. This extension appears to be a common feature in insect flight muscles. (Bullard *et al.,* 1988, Peckham *et al.,* 1992, Qiu *et al.,* 2003, Agianian *et al.,*2004)

1.3 Troponin C structural analysis.

The primary sequence and three-dimensional structure of troponin C are well conserved across vertebrates and invertebrates. All known TnCs share the same fold. They are all formed from two non interacting domains connected by a linker. The way the two domains are connected has attracted much debate. In some of the crystal structures, the

linker between the two domains is a straight helix which confers a dumbbell shape to the protein (Herzberg and James, 1985; Houdusse *et al.,* 1997; Satyshur *et al.,* 1988), whereas in solution the linker is unstructured so that the relative orientation of the two domains is highly underdetermined. (Blumenschein *et al.,* 2005; Dvoretsky *et al.,* 2002; Slupsky and Sykes, 1995; Sia *et al.,* 1997; Vinogradova *et al.,* 2005)

Each domain in TnC contains two EF hands that are connected by a short (2-3 residues) anti-parallel β-sheet. An EF hand is a common motif that consists of two α-helices flanking a 12 residues long loop which can contain the six residues that coordinate the Ca^{2+} ion. In general an EF hand containing protein is defined as being in its holo form when calcium loaded and in its apo form when calcium free. The coordination geometry of the calcium ion is pentagonal bipyramidal.

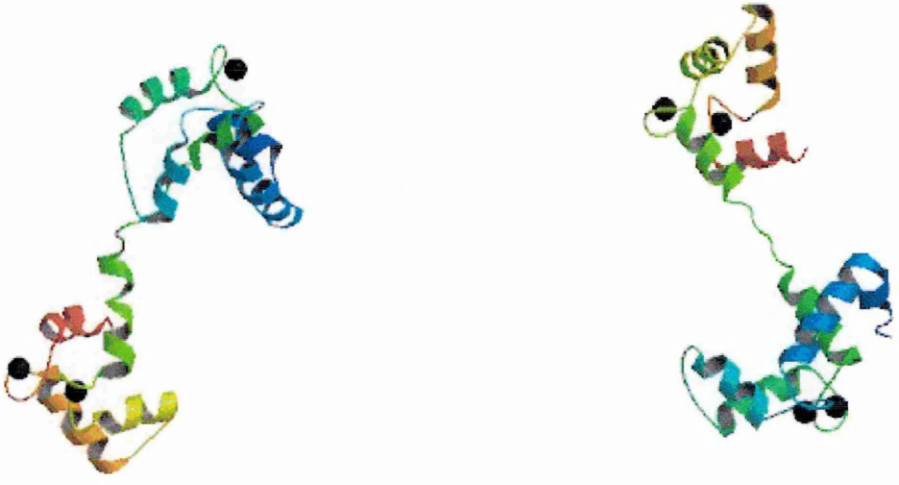


Fig.1.3.1 Ribbon representation of the cardiac (left) (PDB entry: 1aj4) and skeletal (right) (PDB entry: 5tnc) calcium-loaded TnC structures. The calcium ions are shown as black spheres. The skeletal isoform binds four calcium ions, one in each of the EF hand motifs. The cardiac isoform binds only three calcium ions: the first EF-hand motif in the N-terminal domain does not bind calcium. The linker between the N and C-terminal domains is unstructured.

EF-loop position	1	2	3	4	5	6	7	8	9	10	11	12
coordinating ligand	X sc		Y sc		Z sc		-Y bb		-X sc*			-Z sc2
most common	Asp 100%	Lys 29%	Asp 76%	Gly 56%	Asp 52%	Gly 96%	Thr 23%	Ile 68%	Asp 32%	Phe 23%	Glu 29%	Glu 92%
also frequently observed		Ala Gln Thr Val Ile Ser Glu Arg	Asn	Lys Arg Asn	Ser Asn		Phe Lys Gln Tyr Glu Arg	Val Leu	Ser Thr Glu Asn Gly Gln	Tyr Ala Thr Leu Glu Lys	Asp Lys Ala Pro Asn	Asp

Table 1.3.1 The table above, taken from the recently published review by Gifford *et al.* 2007, shows the sequence consensus for an EF-hand loop. The second row shows the calcium coordinating ligands and whether they coordinate the ion through the side-chain (sc) or the backbone (bb). The percentage occurrence of the most common residues is shown in the third row and other residues frequently observed are listed in the fourth row.

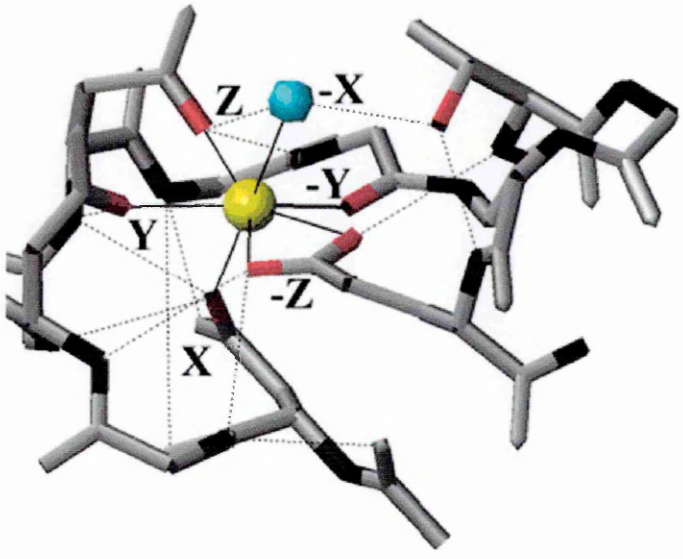


Fig. 1.3.2 The pentagonal bipyramidal coordination geometry for the calcium ion is shown. The – X coordination position is through a water molecule (shown is blue).

The effects of calcium binding on the three-dimensional structure of an EF-hand protein vary remarkably. For example in the case of the C-terminal domain of skeletal TnC calcium binding causes the C-terminal domain to gain structure. In fact, this domain is unstructured in the apo (calcium free) form. (Mercier *et al.*, 2000) In other cases, such as the N-terminal domain of skeletal TnC and calmodulin, the apo form is already folded; ion binding in these cases leads to a conformational change through which the protein is able to transduce the message of an increased Ca^{2+}concentration. (Zhang *et al.*, 1995) The conformational change is usually described as a transition from a closed state to an open state. (Slupsky and Sykes, 1995) The terms "closed" and "open" refer to the arrangement of the helices, the inter-helical angles in the apo state are ~130°-140° whereas in the calcium loaded (holo) state are ~90°. The transition from a closed to an open state causes the exposure of the hydrophobic core which targets can bind. (Mercier *et al.*, 2001)
Finally in some cases, such as the N-terminal domain of cardiac TnC, calcium binding does not cause any major structural rearrangement and the inter-helical angles remain roughly the same. (Spyracopoulous *et al.*, 1998)

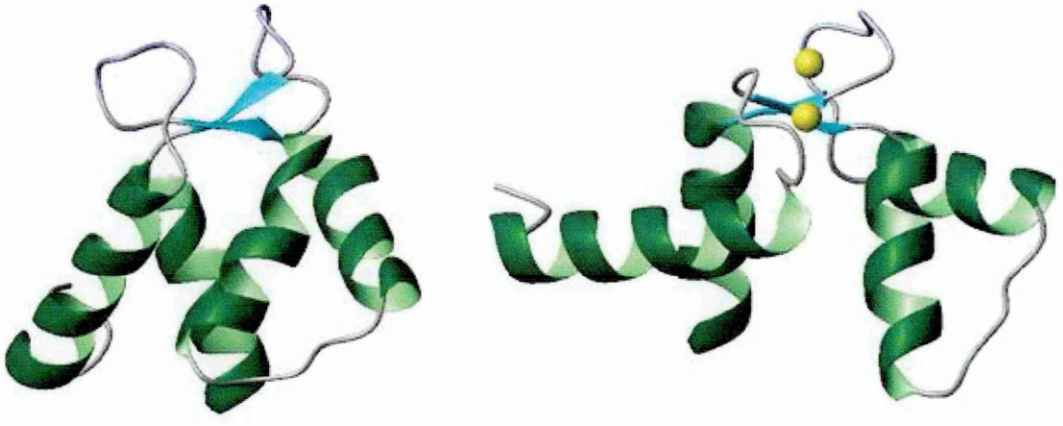


Fig.1.3.3 The transition from a "closed" to an "open" conformation upon calcium binding is shown. The figure shows a calmodulin structure. (pdb code:1cf,1upc) The helices move from a near parallel arrangement to an almost perpendicular one. This rearrangement results in exposure of the hydrophobic core.

1.4 Muscle contraction in vertebrates

Contraction in vertebrates is regulated by calcium. Contraction happens when calcium is released from the sarcoplasmic reticulum (SR) into the myofibrils in response to a motoneuron spike. The generally accepted model to explain calcium regulated contraction is called the "steric blocking model". (Vibert *et al.*, 1997, Gordon *et al.*, 2000, Mak and Smillie, 1981, Potter *et al.*, 1995)

According to this model, tropomyosin blocks access to the myosin binding site of actin in the absence of calcium ions. When the calcium concentration rises, TnC binds to calcium, changes conformation, and causes a structural rearrangement of tropomyosin and actin which makes the myosin binding site on actin accessible. Now the ATP cycle and cross bridges formation can happen causing the thin and thick filament to slide.

A more detailed description of this mechanism in vertebrates is that the C-terminal domain of TnC is believed to be continuously occupied with metal ions. The C-terminal domain is bound to the N-terminal amphiphilic α-helix of TnI which anchors the full TnC to the rest of troponin. At low Ca^{2+} concentrations (i.e. in the relaxed state), the TnI inhibitory region (residues 137–148 in human cardiac TnI) interacts with actin, and thereby inhibits the actin-myosin interaction. The increase in intracellular Ca^{2+} concentration (i.e. the contraction state) causes Ca^{2+} binding to the TnC N-terminal domain which in turn binds to the second amphiphilic α-helix of TnI, that is immediately downstream of the inhibitory region. The inhibitory region is then detached from actin, resulting in the release of inhibition. Myosin can interact with actin and the contraction cycle can start. The X-ray structure of cardiac and skeletal troponin trimer seems to confirm this type of mechanism. (Takeda *et al.*, 2003, Vinogradova *et al.*, 2005).

The entire process of muscle contraction is extremely fast. All the processes involved happen over a very short time period. Calcium binding is essentially diffusion limited, and the electrochemical gradient driving calcium out of the SR into the myofibrils is enormous. (Dickinson, 2006) In contrast, muscle relaxation is a slow process because it requires pumping of calcium back into the SR against the electrochemical gradient. The long time needed to pump back calcium into the SR from the myofibrils limits the frequency rate of contraction in calcium regulated muscles. The slow deactivation process is also the reason why the power output of conventional vertebrate muscles

deteriorates at frequencies well below those used by the wing muscles of small and medium sized insects.

In fact a calcium regulated muscle needs a big SR to increase the frequency of contraction and accelerate the rate of pumping calcium out of the myofibrils. The bigger is the SR the quicker is the pumping out of calcium from the myofibrils, and hence the deactivation process is shorter allowing a higher contraction frequency. Anyhow the hypertrophy of the SR comes at expenses of contractile filaments (i.e. sarcomere units) creating a fundamental trade off between deactivation speed and power output. A typical example of such phenomenon are the very fast oscillating muscles of rattlesnake rattles. Such muscles have a small power output and a high frequency of contraction. (Dickinson, 2006)

1.5 Muscle contraction in insects

A common features of many insect species is the ability to contract wing muscles at a very high frequency. The power output of these muscles does not deteriorate at higher frequencies. In insects the muscles responsible for this feature are called indirect flight muscles. (Pringle JW 1949 and 1978) These muscles, alongside the canonical calcium regulated mechanism, can also contract in an asynchronous manner. The term asynchronous comes from the fact that the frequency of contraction is not coupled with the pre-synaptic motor neuron spikes and therefore is not coupled with the frequency of pumping in and out of calcium from the SR. In the asynchronous model the regulation of muscle contraction is somehow achieved through mechanical stretch, nevertheless calcium has a role also in this second mechanism. Many experiments in fact show that IFM are unable to contract in either way (i.e. calcium regulated and stretch activated) in the complete absence of calcium. (Gordon and Dickinson, 2006, Agianian *et al.*, 2004) This is the reason why the motor neurons of IFM muscles fire continuously at a rate that, even if not coupled to the contraction frequency, is enough to keep the calcium concentration in the sarcomere above the threshold necessary for stretch activation to happen.

In such mechanism, deactivation is not anymore rate limited by the pumping back of calcium into the SR. This is why IFM muscles do not need a big SR and can produce a higher power output at high frequencies.

Stretch activation in IFM of many insect species has a regulatory role and is crucial for the flight ability. It is less known however that stretch activation is a common feature to all muscle fibres: All muscle fibres with a sarcomere architecture will exhibit a certain degree of stretch activation. The degree of tension generated upon stretch will vary according to the type of muscle fibre. Muscles such as skeletal and cardiac exhibit a less degree of stretch activation than IFM. Many models have been proposed to explain such a phenomenon. Some researchers state that a stretched sarcomere reduces the distance between the thin and thick filament increasing the calcium sensitivity and the probability that cross bridges can form. (Wray J.S., 1979)

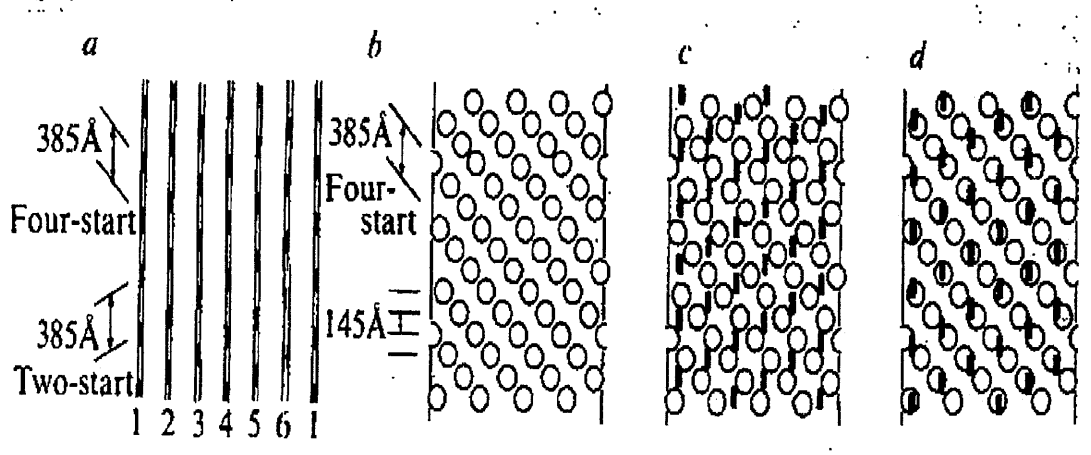


Fig.1.5.1 The figure, taken from Wray J.S 1979 paper, shows the Wray hypothesis to explain stretch activation: a) is the radial projection of the actin envelope, the black bars represent regions of the thin filament accessible to cross bridge attachment b) is the myosin surface lattice where the circle represent cross bridges array; c) and d) show the effect of stretching on the muscle fibres. Due to the stretching the interacting site on the thin and thick filament overlap.

Other models instead focus on the role of the giant cytoskeletal protein titin. Titin length spans half of the sarcomere from the Z disk to the M-line and therefore is well positioned to act as a stretch sensor, furthermore titin is known to have actin and myosin binding sites along its sequence (Li *et al.,* 1995). A third model is based on the fact that TnC alone is sufficient to cause the different stretch induced behaviour in different types of muscles. This hypothesis was first proposed by Gulati and co-workers. (Gulati *et al.,* 1991) They showed that exchanging either fast skeletal TnC isoform in cardiac muscles or vice-versa in skinned muscle fibres was sufficient to cause the corresponding stretch activated behaviour, the same hypothesis was later confirmed by Bullard and co-workers, this time using IFM fibres from *Lethocerus.* (Qiu *et al.,* 2003, Agianian *et al.*, 2004)

The Gulati and co-workers hypothesis is the working hypothesis on which this thesis is based. This hypothesis is also supported by Bullard and co-workers recent discovery that shows how the type of muscle contraction regulation in IFM is determined by the type of TnC isoform present in the troponin trimer. Bullard and co-workers used IFM of *Lethocerus*, a giant water bug, as a working model. They chose *Lethocerus* because its IFM fibres are longer than those of smaller insects making them easier to handle during experiments.

They discovered that in IFM fibres two different types of troponin trimers are present. The two isoforms are identical except for the TnC subunit. The first isoform contains a TnC (F1) subunit that is able to bind only one calcium in its C-terminal domain whereas the second isoform contains a different TnC (F2) subunit able to bind two calcium ions (one in each domain). Using skinned IFM fibres, they demonstrated that the type of regulation depends on which of the two TnC subunits is present in the troponin trimer. Fibres in which all the F2 isoforms were substituted by the F1 isoform were stretch activated, whereas fibres, in which all the F1 isoforms were substituted with F2, had a calcium regulated contraction mechanism. (Agianian *et al.*, 2004)

1.6 Stretch activation, what is it and why study it?

Stretch activation can be defined as the ability of a muscle to produce a delayed force when stretched. When a muscle is stretched to a new isometric length there is a multiphase response, the first step is an increase in the force coincident with the stretch.

Then the force rapidly decays near to the pre-stretch levels and finally there is a delayed increase in force that leads to a new steady state. That second delayed force production is defined as the stretch activated muscle contraction and the delay and the amplitude of the force produced vary according to the type of muscle fibres.(Davis and Roger, 1995, Lombardi *et al.,* 1995, Piazzesi *et al.,*1997, Dobbie *et al.,* 1998) In spite of the fact that most of the current models do not take it into account, stretch activation is a feature common to all types of muscle. Heart muscles and to a smaller extent skeletal muscles exhibit stretch activation. For instance the ability of the heart muscle to be stretch activated has been known since the nineteenth century. Frank and Starling were the first to observe in dog heart muscles that the more the ventricle is filled with blood during diastole, the greater the volume of ejected blood will be during the resulting systolic contraction. (Frank, 1895, Starling and Visscher 1926) This means that the force of contraction will increase as the heart is filled with more blood and the increased force produced is a direct consequence of the effect of a bigger load on a single muscle fibre. Increased load further stretches the myocardium and enhances the muscle fibres power output. Based on these observations in 1914 Ernest Starling hypothesized that "the mechanical energy set free in the passage from the resting to the active state is a function of the length of the fibre." (Starling, 1918)

As an empirical observation, the stretch of a muscle fibre can increase its output and the extent of this increase depends on the type of muscle fibre and the range of frequencies at which the muscle works.

1.7 Aims of this thesis

The aim of this thesis has been to understand the molecular basis of stretch activation and in particular the role of calcium in stretch activation. In the first part the solution structure of holo (calcium-bound) and apo (calcium-free) F1 and the comparison to the other known isoforms will be described. In the second part the interaction between F1 and the rest of the troponin trimer subunits, through the use of synthetic peptides and recombinant technique, is analyzed. The possible role of calcium in the regulation of these interactions will also be discussed. Finally, a model to explain stretch activated muscle contraction based on the findings reported in this thesis is proposed.

CHAPTER 2

BIOPHYSICAL TECNIQUES

2.1 Fluorescence spectroscopy

Fluorescence is a spectroscopic technique that is often used as a tool to monitor protein ligand interactions. (Creighton T.E, Freeman& Co 1993) It is an emissive phenomenon based on the emission between two different electronic states of a fluorophore. It involves using a beam of light to excite the electrons of a certain fluorophore that in turn will relax and emit light at a lower energy (i.e. longer wavelength). The aromatic residues Tyr and Trp are the major contributors to the intrinsic fluorescence of proteins. (Creighton T.E, Freeman& Co 1993) The emission maximum for Tyr is at around 306 nm. Changes in the environment around Tyr in proteins sometimes cause a change in the quantum yield but the emission maximum generally stays close to 306 nm. In the case of Trp both the quantum yield and the emission maximum depend strongly upon environment. A completely solvent exposed Trp has an emission maximum at around 355 nm; more buried Trp residues have a blue shifted (i.e. shorter wavelength) emission maxima which may be as short as 322 nm in extreme cases. (Cantor and Schimmel, Wiley & Sons 1980) The extent of blue shift is related to the hydrophobicity of the environment.

The small amount of sample required and the short data collection time make fluorescence a useful tool in the study of protein ligand interactions and protein stability.

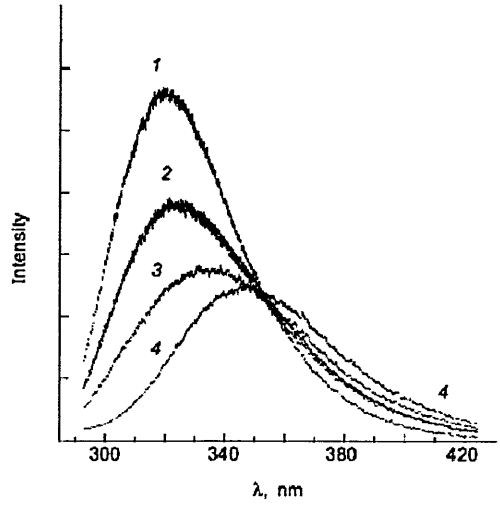


Fig. 2.1.1 Typical fluorescence spectra of a Trp containing protein. The curves marked 1 to 4 were acquired at different concentration of denaturant. Curve 1 is with no urea, curve 4 is with 5M urea. Denaturation of the protein causes the exposure of the Trp residue which leads to changes in the fluorescence emission spectrum

2.1.1 Direct fluorescence titrations: theory and data analysis

In this paragraph I shall show how to interpret a direct fluorescence titration data between a fluorescent peptide and a non fluorescent protein F1, since this analysis will be used later on in this thesis. The fluorescent peptide used here is melittin (M) the principal active component of bee venom.(Ficher and Neumann 1961) It was chosen because is known to bind to F1 and EF hand protein in general with a 1:1 stoichiometry (Comte, M *et al.,* 1983) as expressed by the below equation:

$$F1 + M \leftrightarrow F1M \qquad \text{with } K_d = [F1]\,[M] / [F1M] \tag{1}$$

$F1_{tot}$ and M_{tot} are respectively the total concentration of F1 and melittin and they are given by the below relations:

$$F1_{tot} = [F1] + [F1M] \tag{2}$$

$$M_{tot} = [M] + [F1M] \tag{3}$$

Therefore:

$$K_d = (F1_{tot} - [F1M])\,(M_{tot} - [F1M]) / [F1M] \tag{4}$$

[F1M] is the root of the quadratic:

$$[F1M]^2 - [F1M](K_d + F1_{tot} + M_{tot}) + (F1_{tot}\ M_{tot}) = 0 \tag{5}$$

The remaining concentration can then be calculated using:

$$[M] = M_{tot} - [F1M] \quad \text{and} \quad [F1] = F1_{tot} - [F1M] \tag{6}$$

The fluorescence signal (S) measure is:

$$S = S_{(F1)}[F1] + S_{(M)}[M] + S_{(F1M)}[F1M] \quad (7)$$

where $S_{(F1)}$, $S_{(M)}$ and $S_{(F1M)}$ are the specific fluorescence intensities of F1, M and F1M respectively.

The data were fitted to the above equation with the specific fluorescence intensities and K_d as variables. A standard non linear least squares fitting procedure written by Dr. Stephen Martin (N.I.M.R.) was used.

2.1.2 Alternative data analysis for direct fluorescence titrations

In this paragraph I am going to show a different way of analyzing direct fluorescence titration data, that does not require the solving of a quadratic equation nor any fitting procedure. At the mid point of the titration (50% of the total fluorescence change completed) the concentration of the complex ([F1M]) is by definition equal to the concentration of uncomplexed melittin ([M]) and therefore equation (1) can be written:

$$K_d = [F1] \quad (8)$$

That is: the concentration of uncomplexed F1 at the mid point ([F1]) is equal to the K_d. The concentration of uncomplexed F1 at the mid point can easily calculated as the total concentration of F1 added at the mid point minus one half of the initial melittin concentration.

2.1.3 Fluorescence competition assays: theory and data analysis

Competition assays can be used to measure binding affinities for interaction that are not accompanied by any significant change in an optical signal. This approach was used here to measure the binding of F1 to TnH peptides: neither F1 nor the peptides has a suitable fluorescence signal. The basic idea behind this method is to measure the binding affinity between F1 and a fluorescent molecule in this case melittin (M) and then to measure how the binding of melittin is affected by the presence of the TnH peptides. Below is reported a way to analyze such fluorescence competition assay data. F1 interacts with melittin and TnH according to the following scheme:

$F1 + M \leftrightarrow F1M$ with $K_M = [F1M] / ([F1] [M])$ (1)

$F1 + TnH \leftrightarrow F1TnH$ with $K_{TnH} = [F1TnH] / ([F1] [TnH])$ (2)

$F1_{tot}$, M_{tot} and TnH_{tot} are the total concentration of F1, M and TnH and are given by:

$F1_{tot} = [F1] + [F1M] + [F1TnH]$ (3)

$M_{tot} = [M] + [F1M]$ (4)

$TnH_{tot} = [TnH] + [F1TnH]$ (5)

Then:

$[M] = [F1M] / (K_M [F1])$ $M_{tot} = ([F1M] / K_M[F1]) + [F1M]$ (6)

$M_{tot} = [F1M] (1+(K_M[F1])^{-1}) = [F1M] (K_M[F1] + 1) / K_M[F1]$ (7)

$[F1M] = M_{tot}K_M[F1] / (K_M[F1] +1)$ (8)

In the same way it can be shown that:

$[F1TnH] = TnH_{tot} K_{TnH} [F1] / (K_{TnH}[F1] + 1)$ (9)

Substituting these expression for [F1M] and [F1TnH] in the expression for $F1_{tot}$ it gives:

$F1_{tot} = [F1] + (F1_{tot}K_M[F1]) / (K_M[F1] +1) + (TnH_{tot} K_{TnH}[F1])/ (K_{TnH}[F1]+1)$ (10)

Expanding equation (10) gives an expression for [F1] as the root of the cubic equation:

$[F1]^3 (K_{TnH}K_M)+[F1]^2(-F1_{tot}K_{TnH}K_M+ K_M+ K_{TnH} +M_{tot}K_{TnH}K_M +TnH_{tot}K_{TnH}K_M) +$

$TnH_{tot}K_{TnH}K_M) + [F1] (-F1_{tot}K_M - F1_{tot}K_{TnH} + 1 +M_{tot}K_M + TnH_{tot}K_{TnH}) - F1_{tot} = 0$ (11)

And the remaining concentration can be calculated as follows:

$$[F1M] = (M_{tot}K_M[F1])/(K_M[F1]+1) \quad (12)$$

$$[F1TnH] = (TnH_{tot}K_{TnH}[F1])/(K_{TnH}[F1]+1) \quad (13)$$

$$[M]=[F1M]/K_M[F1] \quad (14)$$

$$[TnH] = [F1TnH]/K_{TnH}[F1] \quad (15)$$

The fluorescence intensity (S) is given by:

$$S = S_{(F1)}[F1]+S_{(M)}[M]+S_{(F1TnH)}[F1TnH]+S_{(F1M)}[F1M] \quad (16)$$

Where $S_{(F1)}$, $S_{(M)}$, $S_{(F1TnH)}$ and $S_{(F1M)}$ are the specific fluorescence intensities of F1, M, F1M and F1TnH respectively. The experimental data were fitted to the above equation with the specific fluorescence intensities and the unknown association constant K_{TnH} as variables and K_M instead was held constant at the value measured with the direct fluorescence titration. A standard non linear least square procedure written by Dr Stephen Martin (N.I.M.R.) was used in the analysis.

2.1.4 Alternative data analysis for fluorescence competition assays

In this paragraph I am going to explain a different way to analyze the competition titration data that does not require the solving of a cubic equation. The K_d values obtained through this different fitting procedure are similar to the one obtained using the standard fitting.

Two limiting cases can be considered. The first case is where the silent (non-fluorescent) binding peptide binds to F1 much more strongly than melittin does, the second case is where the silent peptide binds less strongly to F1 than melittin does. The titrations have to

be done in different ways. In the first case the F1 is first saturated with melittin (at that point (A) the total concentrations of F1, M and F1M can be calculated). The melittin is then displaced by adding increasing quantities of the silent peptide. When half of the total signal change has occurred the F1M concentration must be half the concentration at point A.

Because K_M is known it is possible to calculate [F1] from the expression for K_M:

$$F1 + M \leftrightarrow F1M \qquad \text{with } K_M = [F1M] / ([F1]\,[M]) \qquad (1)$$

And since

$$F1 + TnH \leftrightarrow F1TnH \qquad \text{with } K_{TnH} = [F1TnH] / [[F1]\,[TnH] \qquad (2)$$

In fact is possible to calculate all the equilibrium concentrations

$$[F1] = K_M\,[F1M] / [M] \qquad (3)$$

$$[F1TnH] = F1_{tot} - [F1] - [F1M] \qquad (4)$$

$$[TnH] = TnH_{tot} - [F1TnH] \qquad (5)$$

In the second case instead the melittin is titrated with a stock solution of F1 in the presence of different large excess concentrations of the silent peptide.

Equations (1) and (2) still apply but is possible to write a third equation (K_{obs}) derived from equation (1) where the total F1 concentration is the sum of free and bound F1

$$K_{obs} = [M]([F1] + [F1TnH]) / [F1M] \qquad (6)$$

Then substituting [F1TnH] from equation (2)

$$K_{obs} = ([M]([F1] + ([F1][TnH])/K_{TnH}))/ [F1M] \qquad (7)$$

And

$$K_{obs} = [M]\,[F1]\,(1+([TnH]/K_{TnH}))\,/\,[F1M] \quad (8)$$

Introducing K_M from equation (1)

$$K_{obs} = K_M\,(1+([TnH]/K_{TnH})) \quad (9)$$

Then because there is an excess of TnH it is possible to assume that [TnH] is equal to TnH_{tot}. Finally is possible to calculate K_{TnH} once K_{obs} is obtained through the fitting as for a direct fluorescence titration.

2.2 CD spectroscopy

CD is an absorptive phenomenon generated by the differential absorption of left- and right-handed circularly polarized light. A circularly polarized wave can be seen as the superimposition of two waves plane-polarized in two perpendicular planes that have the same amplitude and wavelength and that have a phase difference of 90 degrees. The phase difference, either 90 degrees or -90 degrees, gives rise to a right and left circularly polarized wave. When such a wave goes through an asymmetric medium the left and right component are absorbed differently and cause a CD signal. In the case of proteins the "asymmetric" absorptive unit is the peptide bond that exhibits different differential absorption properties depending on the secondary structure element it is in. The peptide bond generates two transitions: n-π^* (forbidden) and π-π^* (allowed), where n, π, π^* are the non-bonded, bonded and anti-bond orbitals in a peptide bond according to the molecular orbital theory. The first transition involves non-bonding electrons of the carbonyl oxygen and is centred around 220nm, the second one involves π electrons of the carbonyl and is centred around 190nm. (Cantor and Schimmel, Wiley & Sons 1980) These wavelength values change depending on which secondary structure the peptide bond is in. In an α-helix structure three bands are present: a positive one for n-π^* at 190-195 nm, and two negatives ones for π-π^* at 208nm and n-π^* at 222nm (Cantor and

Schimmel, Wiley & Sons 1980) In a β-sheet structure there is a positive band for π-π* at 195-200 nm and a negative one for n-π* at 215-220nm and in a random structure there is a negative band for the π-π* at 200nm and a positive or negative one for n-π* at 220nm.

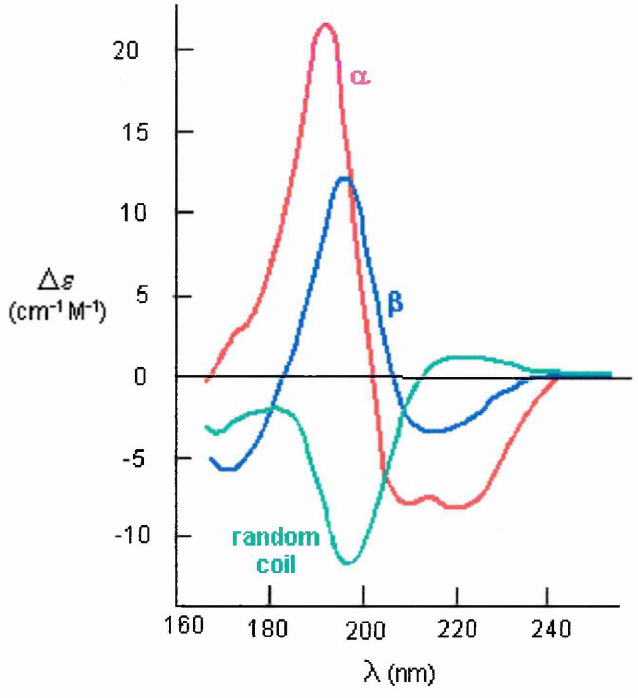


Fig.2.2.1 The CD Spectra of alpha-Helix, beta-Sheet, and random Coils

Several fitting algorithms can be used to calculate the percentage of each secondary structure element present in a protein. Fitting procedures based on reference spectra of protein with known X-ray structure, give reasonable estimate of the secondary structure content. All these methods assume that the CD spectrum in a protein is the linear combination of α, β, coil and random structures, that the effect of tertiary structure on CD is negligible, that the ensemble averaged solution structure is equivalent to the X-ray structure, and that the geometric variability of secondary structures need not be explicitly considered. Whereas the first two assumptions can be considered to be true, the last assumption is in general not true and this in part may account for the differences between X-ray measured secondary structure and CD one.

In addition, CD is a convenient method to study protein unfolding transition and to measure the thermodynamic stability of proteins. Thermal and chemical denaturation curves can be obtained by monitoring the change of ellipticity at a fixed wavelength. These curves can be used to derive equilibrium constants, free energies of unfolding and the mid-point temperature for thermal unfolding when analyzed with the appropriate model. A qualitative analysis of the cooperativity of the unfolding curves can also give some information about the kinetics of the protein in solution. The cooperativity of the unfolding reaction is measured by the width and shape of the unfolding transition. A very steep transition reaction indicates a highly cooperative unfolding where the protein existed initially as a compact, well-folded structure. A broad transition indicates a very gradual, non-cooperative melting reaction where the protein existed initially as a very flexible, partially unfolded protein or as a heterogeneous population of folded structures.

2.2.1 Thermal unfolding: theory and data analysis

Thermal unfolding of a protein such us F1 with two domains (N and C) that tumble almost independently in solution can be described, to a good approximation, as the unfolding of a mixture of two separate proteins A and B.

The ΔG values for the two transitions are given by the appropriate Gibbs-Helmholtz equations:

$$\Delta G_A = \Delta H_{m,A}(1-T/T_{m,A}) + \Delta C_{p,A}(T-T_{m,A}-T\ln(T/T_{m,A}) \quad (1)$$

$$\Delta G_B = \Delta H_{m,B}(1-T/T_{m,B}) + \Delta C_{p,B}(T-T_{m,B}-T\ln(T/T_{m,B}) \quad (2)$$

where T, T_m, ΔH_m and ΔC_p are the absolute temperature, the transition midpoint temperature, the enthalpies changes at the midpoint and the heat capacity changes respectively.

The optical signals are considered to be related to the temperature in a linearly dependent way:

$$Opt_{N,A} = I_{N,A} + S_{N,A}T \quad (3)$$

$$Opt_{U,A} = I_{U,A} + S_{U,A}T \quad (4)$$

$$Opt_{N,B} = I_{N,B} + S_{N,B}T \quad (5)$$

$$Opt_{N,B} = I_{U,B} + S_{U,B}T \quad (6)$$

Where I and S are the intercepts and the slopes for the native and unfolded CD signal respectively .The change in the optical signal as a function of temperature is given by

$$Signal = (I_{N,A}+S_{N,A}T)\, F_{N,A}+(I_{U,A}+S_{U,A}T)F_{U,A}+ (I_{N,B}+S_{N,B}T)F_{N,B}+(I_{U,B}+S_{U,A}T)F_{U,B} \quad (7)$$

Where

$$F_{U,A} = K_A/(1+K_A) = (e^{-\Delta G°A/RT})/(1+ e^{-\Delta G°A/RT}) \quad (8)$$

and

$$F_{U,B} = K_B/(1+K_B) = (e^{-\Delta G°B/RT})/(1+ e^{-\Delta G°B/RT}) \quad (9)$$

A standard non linear least squares fitting procedure written by Dr. Stephen Martin (N.I.M.R.) was used to fit the data to the above equation with T, T_m, ΔH_m, ΔC_p, I and S as variables.

2.2.2 Chemical unfolding: theory and data analysis

Chemical unfolding of a protein such us F1 with two domains (N and C) that tumble almost independently in solution can be described, to a good approximation, as the unfolding of a mixture of two separate proteins A and B.

The free energies of unfolding for the two proteins (A and B) are assumed to depend on the denaturant concentration (D) according to

$$\Delta G_A = \Delta G°_A + m_A[D] \quad (1)$$

$$\Delta G_B = \Delta G°_B + m_B[D] \quad (2)$$

Where $\Delta G°_A$ and $\Delta G°_B$ are the free energies for unfolding in the absence of the denaturant and m_A and m_B are the constants which describe the linear dependence of the ΔGs on denaturant concentrations. K_A and K_B can be expressed as:

$$K_A = \exp(-\Delta G_A / RT) \quad (3) \quad \text{and} \quad K_B = \exp(-\Delta G_B / RT) \quad (4)$$

And the fraction of native ($F_{N,A}$ and $F_{N,B}$) and unfolded ($F_{U,A}$ and $F_{U,B}$) is expressed as

$$F_{U,A} = K_A/(1+K_A) \quad (5) \quad \text{and} \quad F_{N,A} = 1 - F_{U,A} \quad (6)$$

$$F_{U,B} = K_B/(1+K_B) \quad (7) \quad \text{and} \quad F_{N,B} = 1 - F_{U,B} \quad (8)$$

The optical signals of the native and unfolded protein are assumed to be linearly dependent on the denaturant concentration as follows:

$$Opt_{N,A} = I_{N,A} + S_{N,A}[D] \quad (9)$$

$$Opt_{U,A} = I_{U,A} + S_{U,A}[D] \quad (10)$$

$$Opt_{N,B} = I_{N,B} + S_{N,B}[D] \quad (11)$$

$$Opt_{N,B} = I_{U,B} + S_{U,B}[D] \quad (12)$$

Where I and S are respectively the intercepts and the slopes for the native and unfolded CD signal. The change in optical signal as a function of denaturant concentration can be expressed as follows:

$$Signal = (I_{N,A}+S_{N,A}[D])\, F_{N,A}+(I_{U,A}+S_{U,A}[D])F_{U,A}+ (I_{N,B}+S_{N,B}[D])F_{N,B}+(I_{U,B}+S_{U,A}[D])F_{U,B} \quad (13)$$

Where

$$F_{U,A} = K_A/(1+K_A) = (e^{-\Delta G°A/RT})/(1+ e^{-\Delta G°A/RT}) \quad (14)$$

and

$$F_{U,B} = K_B/(1+K_B) = (e^{-\Delta G°B/RT})/(1+ e^{-\Delta G°B/RT}) \quad (15)$$

A standard non linear least squares fitting procedure written by Dr. Stephen Martin (N.I.M.R.) was used to fit the data to the above equation with $\Delta G°_A$, $\Delta G°_B$, m_A, m_B, I and S as variables.

2.2.3 Effects of ligand binding on chemical unfolding: theory and data analysis

Ligand binding to the native and or unfolded states of a protein affects the stability of the protein itself. In the case of F1 just the C terminal domain of the protein interacts with TnH peptides and the analysis of the urea unfolding data shows that apo F1 undergoes a two-state transition (N and C domains unfolding) at reasonably different denaturant concentrations. Given these condition, it is possible to study the apo F1 chemical denaturation in the presence of TnH as a two-state unfolding system.

In such system the net interaction free energy is the difference in interaction free energy for the native and the unfolded state in the presence and in the absence of the ligand:

$$\Delta\Delta G = \Delta G_{(L)} - \Delta G_{(F)} \quad (1)$$

Where $\Delta G_{(L)}$ and $\Delta G_{(F)}$ are the free energy changes for the unfolding reaction with and without ligand.

In the presence of the ligand and assuming the ligand is binding only to the native form of the protein the system can be expressed in the following way:

$$NL \leftrightarrow L + N \leftrightarrow U \quad (2)$$

Where NL, L, N, U are the complex, the ligand, the native form of the protein and the unfolded form respectively. The following relation can be written:

$$K_u = [U]/[N] \quad (3) \qquad K_b = [NL]/[N]\,[L] \quad (4)$$

Considering the presence of the ligand the new unfolding constant K_l is

$$K_l = [U]/([N]+[NL]) = [U]/([N]+K_b[N][L]) = [U]/([N](1+K_b[L])) = K_u/(1+K_b[L]) \quad (5)$$

Therefore

$$\Delta G_{(L)} = -RT\ln K_l = -RT\ln K_u/(1+K_b[L]) \quad (6)$$

$$\Delta G_{(F)} = -RT\ln K_u \quad (7)$$

$$\Delta\Delta G = \Delta G_{(L)} - \Delta G_{(F)} = -RT\ln K_u/(1+K_b[L]) + RT\ln K_u = -RT\ln K_u + RT\ln(1+K_b[L]) + RT\ln K_u \quad (8)$$

$$\Delta\Delta G = +RT\ln(1+K_b[L]) \quad (9)$$

Assuming that [L] is equal to the total ligand concentration (this is a safe assumption because the experiment is performed is the presence of a large excess of ligand) and measuring $\Delta G_{(L)}$ and $\Delta G_{(F)}$ it is possible to derive the ligand binding constant.

2.3 F1 Calcium affinities: theory and data analysis

The first step is to measure affinity of the chromophoric chelator 5,5' Br_2-BAPTA for calcium. The chelator (C) is known to bind to calcium with a 1:1 stoichiometry and absorbs in the UV region:

$Ca + C \leftrightarrow CaC$ (1) with $K_d = [Ca]\,[C] / [CaC]$ (2)

and the total concentration of Ca and chelator (Ca_{tot} and C_{tot}) are given by

$Ca_{tot} = [Ca] + [CaC]$ (3)

$C_{tot} = [C] + [CaC]$ (4)

Therefore

$K_d = (Ca_{tot} - [CaC])(C_{tot} - [CaC]) / [CaC]$ (5)

[CaC] is the root of the quadratic:

$[CaC]^2 - [CaC](K_d + Ca_{tot} + C_{tot}) + Ca_{tot}C_{tot} = 0$ (6)

and the remaining concentration are calculated using:

$[C] = C_{tot} - [CaC]$ (7)

$[Ca] = Ca_{tot} - [CaC]$ (8)

The measured absorbance at 263nm is given by:

$A = \varepsilon_{(C)}[C] + \varepsilon_{(CaC)}[CaC]$ (9)

where $\varepsilon_{(C)}$ and $\varepsilon_{(CaC)}$ are the extinction coefficient of the free chelator and the complex CaC.

The experimental data were fitted to the above equation with K_d, $\varepsilon_{(C)}$ and $\varepsilon_{(CaC)}$ as variables. A standard non linear squares fitting procedure written by Dr. Stephen Martin (N.I.M.R.) was used. The Br_2-BAPTA calcium association constant ($K_{Ca} = K_d^{-1}$) was measured at 25°C and 40 °C.

The second step of the experiment is to titrate a mixture of cheletor and apo F1with a stock solution of calcium.

The experimental system is expressed by the following equation:

$$F1 + Ca \leftrightarrow F1Ca \qquad K_{F1Ca} = [F1Ca] / ([F1]\,[Ca]) \qquad (10)$$

$$Ca + C \leftrightarrow CaC \qquad K_{CaC} = [CaC] / ([Ca]\,[C]) \qquad (11)$$

Ca_{tot}, C_{tot} and $F1_{tot}$ are the total concentrations of calcium, chelator and F1 and are given by:

$$Ca_{tot} = [Ca] + [CaC] + [CaF1] \qquad (12)$$

$$C_{tot} = [C] + [CaC] \qquad (13)$$

$$F1_{tot} = [F1] + [CaF1] \qquad (14)$$

Then:

$$[C] = [CaC] / (K_{CaC}\,[Ca]) \qquad (15)$$

$$C_{tot} = ([CaC] / K_{CaC}[Ca]) + [CaC] \qquad (16)$$

$$C_{tot} = [CaC]\,(1+(K_{CaC}[Ca])^{-1}) = [CaC]\,((K_{CaC}[Ca] + 1) / K_{CaC}[Ca] \qquad (17)$$

$$[CaC] = C_{tot}K_{CaC}[Ca] / (K_{CaC}[Ca] +1) \qquad (18)$$

In the same way it can be shown that:

$$[CaF1] = F1_{tot} K_{F1Ca} [Ca] / (K_{F1Ca}[Ca] + 1) \quad (19)$$

Substituting these expression for [CaC] and [CaF1] in Ca_{tot} it gives:

$$Ca_{tot} = [Ca] + C_{tot}K_{CaC}[Ca] / (K_{CaC}[Ca] +1) + F1_{tot} K_{F1Ca} [Ca] / (K_{F1Ca}[Ca] + 1) \quad (20)$$

Expanding this equation gives [Ca] as the root of the cubic equation:

$$[Ca]^3(K_{F1Ca}K_{CaC})+[Ca]^2(-Ca_{tot}K_{F1Ca}K_{CaC}+K_{CaC}+K_{F1Ca}+C_{tot}K_{F1Ca}K_{CaC}+F1_{tot}K_{F1Ca}K_{CaC})+$$

$$F1_{tot}K_{F1Ca}K_{CaC} + [Ca] (-Ca_{tot}K_{CaC} - Ca_{tot}K_{F1Ca} + 1 +C_{tot}K_{CaC} + F1_{tot}K_{F1Ca}) - Ca_{tot} = 0 \quad (21)$$

And the remaining concentration can be calculated as follows:

$$[CaC] = (C_{tot}K_{CaC}[Ca])/(K_{CaC}[Ca]+1) \quad (22)$$

$$[CaF1] = (F1_{tot}K_{F1Ca}[Ca])/(K_{F1Ca}[Ca]+1) \quad (23)$$

$$[C]=[CaC]/K_{CaC}[Ca] \quad (24)$$

$$[F1] = [CaF1]/K_{F1Ca}[Ca] \quad (25)$$

The measured absorbance signal (A) is given by:

$$A = \varepsilon_{(C)}[C]+ \varepsilon_{(F1Ca)}[CaF1]+ \varepsilon_{(CaC)}[CaC] \quad (26)$$

Where $\varepsilon_{(C)}$, $\varepsilon_{(CaF1)}$ and $\varepsilon_{(CaC)}$ are the extinction coefficients of C, CaC and CaF1 respectively. The experimental data were fitted to the above equation with the extinction coefficients and the unknown association constant K_{F1Ca} as variables. K_{CaC} was held

constant at the value measured with the direct experiment. A standard non linear least square procedure written by Dr. Stephen Martin (N.I.M.R.) was used.

2.4 NMR spectroscopy

NMR is an acronym for Nuclear Magnetic Resonance. The two basic players in an NMR experiment are: a strong external magnetic field and NMR active nucleus.

The nucleus of atoms is made of elementary particles called neutrons and protons, these particles have an intrinsic property called spin. The overall spin of a nucleus is defined by the spin quantum number I. Nucleus with even number of neutron and protons have a I=0 and therefore they do not have any spin and they do not give rise to any NMR signal, whereas all isotopes with I≠0 can theoretically produce and NMR signal. When a nucleus with a I≠0 interacts with an external magnetic field the nuclear ground state will split in N energy levels according to the following relation N=2I+1. In bio-molecular NMR the more commonly analyzed nuclei are ^{1}H,^{15}N,^{13}C , they all have I =1/2. When a nucleus with I = 1/2 is inserted into a strong magnetic field the nuclear spin becomes aligned, for a nucleus with I =1/2 only two possible orientations with respect to external magnetic field are possible: a parallel one and an anti-parallel one. These two orientation according to the N=2I+1 relation give rise to two energy levels that can be perturbed if irradiated with an appropriate radiofrequency. Such radiofrequencies cause a perturbation of the population of the two states if the irradiating frequency is equal to the frequency difference between the two states. Once the perturbation is stopped the system will then relax and go back to the pre-perturbed state producing a signal that is then detected as an electrical current. The fact that the proton, nitrogen and carbon NMR active nuclei are all I=1/2 is a lucky coincidence. For such nuclei only two energy levels are possible when they interact with an external magnetic field, this feature makes the interpretation of nitrogen, carbon and proton NMR spectra of bio-molecular molecules easier than for example for the NMR active oxygen isotope ^{17}O that has a nuclear spin of I=5/2. In the oxygen case the possible energy levels when interacting with an external magnetic field are six, this feature makes the interpretation of ^{17}O NMR spectra more complicated.

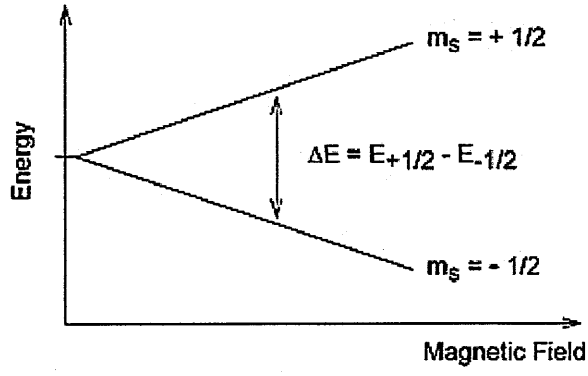


Fig 2.4.1 Behaviour of a I=1/2 nucleus when interacting with an external magnetic field. The stronger is the external magnetic field the bigger is the difference between the two energy levels. Higher energy difference will cause a higher difference in population of the states and therefore a stronger signal.

A comprehensive description of NMR spectroscopy can be found in *Spin dynamics* by Malcolm Levitt (Wiley & Sons 2001) and in *Understanding NMR spectroscopy* by James Keeler (Wiley & Sons 2005). Here I am going to briefly summarize some of the tools used in this thesis.

2.4.1 NMR structure calculation

Through NMR is possible to calculate protein structures at atomic resolution. The strategy generally used can be divided into three steps:

a) *Assignment of the frequency resonances of the atoms belonging to the proteins.* The use of uniformly ^{15}N and ^{13}C labelled recombinant proteins coupled with the use of heteronuclear triple resonance experiments has made the assignment of relatively large proteins much easier for several reasons. First, the fact that these experiments rely exclusively on $^{1}J/^{2}J$ couplings which are relatively large (tens of Hz) if compared to proton-proton higher order J couplings (few Hz) allows a faster magnetization transfer that can compete with the loss of transverse magnetization. This makes it possible to correlate the resonances of different atoms within the protein sequence. Then the use of a combination of different types of heteronuclear triple experiments which correlate different atoms eases the assignment task.

A second big advantage is the increased frequency resolution owing to the higher spectral width for the carbon (200 ppm) and the nitrogen (60 ppm) if compared to the proton (12 ppm). The ^{15}N and ^{13}C chemical shift values themselves add information on the system studied: for instance the carbon chemical shifts are very useful in the assignment of the amino-acid type (i.e. Ser and Thr are the only amino-acids where the Cβ chemical shift value is higher than the Cα , Ala has a very high shifted Cβ, Ile,Pro,Thr,Val in general have a very down shifted value for the Cα) but they can also give information about the secondary structure a residue is in (i.e. α, β and C' carbon are very sensitive to the torsion angle φ and allow one to distinguish between α helix and β-sheet secondary structures).

b) *Collection of distance restraints.* The main source of geometric information relies on the proton-proton distance restrains derived from the nuclear Overhauser effect (NOE). Historically the nuclear Overhauser effect term is used to refer to the steady-state enhancement of one spin species when the population of a second spin species is saturated by a resonant radiofrequency. The saturation causes exchange of z-magnetization between spins. In NOESY experiments no saturation is involved. During the mixing time magnetization is in fact kept along the z- axis and spins are allowed to cross-relax thereby exchanging z-magnetization without any saturation being involved. Two or more spins on the way to thermal equilibrium can exchange z-magnetization by means of the dipolar interaction. The probability for the spins to exchange magnetization and therefore to cross relax is given as a rate constant defined as the NOE build-up rate (K_{NOE}). K_{NOE} is proportional to the square of the dipolar interaction times a function that depends on the rotational correlation time τ_c. For NOE mixing times short compared to the K_{NOE} the increase in z magnetization on spin S due to NOE transfer from spin I is proportional to $\gamma_I^2\gamma_S^2/r^6$ where γ is the gyromagnetic ratio of the spin of interest and r is the distance between the two. It follows that the NOE is strongest between protons (higher γ values) and is inversely proportional to the sixth power of the distance. The NOE effect as defined above is used in NOESY experiments which is one of the most frequently used tools for determining NMR solution structures. Acquiring NOESY experiments allows one to estimate simultaneously a large number of intermolecular distances and thus get a three-dimensional structure.

c) *Solution structure calculation.* The final aim of NMR structure calculation is to build a model that simultaneously satisfies the experimental data (i.e. NOEs, chemical shifts, J coupling constants) and the chemical information (i.e. stereochemistry and non-bonded interactions). This task can be expressed as the search of a global minimum for an hybrid-energy-function. Such an energy function can be expressed as:

$$E = E_{chem} + E_{nmr}$$

The first term describes the agreement with expected values for bond lengths, bond angles, planarity, chirality and non-bonded interactions (van der Waals, hydrogen bonds and electrostatic contributions) whereas the second one describes the agreement with the experimental data.

Many energy terms contribute to the energy function and this causes two main problems that arise when trying to calculate a 3D structure using NMR based data. The first problem is that the energy function has many local minima, so that there is always the possibility of being trapped in one of these during the structure calculation.

Simulated annealing is the procedure usually used to overcome this problem. Annealing is a term which refers to the heating of a solid up to the transition to the liquid phase where the particles are free to rearrange and find the lowest energy state while the system is slowly cooled. Something conceptually very similar is applied in simulated annealing protocols, where the heating is achieved by increasing the particles' velocities. It is worth noting that in this case temperature has no physical meaning but just refers to the likelihood of the system overcoming the barriers of the energy function.

The second problem relates to the ratio between experimental data and variables that, especially with longer protein sequences, may not be big enough for the system to converge. Torsion angle dynamics, based on the observation that deviations from ideal bond lengths and bond angles are usually small, imposes fixed lengths and angles, thus producing a ~10 fold reduction of the number of variables and increasing the ratio experimental data to variables.

Fixed length and angle constraints are particularly useful when using simulated annealing where at high temperatures the conventional molecular dynamic allows for substantial distortion from the ideal geometry.

2.4.2 HSQC

HSQC is an acronym for heteronuclear single quantum coherence. (Norwood *et al.*, 1989) It is a bi-dimensional experiment used to obtain a correlation between the ^{1}H nucleus and the ^{13}C or ^{15}N nuclei. The spectrum contains a peak for each proton bound to a ^{13}C or ^{15}N depending on the type of labelling. The coordinates of the peak correspond to the proton and heteroatom resonance frequency. The type of residue, the three-dimensional arrangement and the possible proximity of an aromatic or unsaturated group will influence the resonance frequencies of the specific signal.

$\{^{1}H, ^{15}N\}$ HSQC is in general preferred to $\{^{1}H, ^{13}C\}$ HSQC when monitoring protein binding studies because of the greater resolution, owing to the smaller number of peaks. The ^{1}H-^{15}N chemical shifts of the backbone moieties are very sensitive to small changes in the backbone conformation (ψ and φ angles), to changes in the anisotropic shielding caused by the proximity of aromatic or unsaturated groups and to the hydrogen bonding pattern. Changes in some or all of the afore mentioned features will usually occur upon ligand binding. This will cause chemical shift changes of backbone signals, allowing one to monitor protein binding. If the ^{1}H-^{15}N chemical shift assignment and the protein structure are known the chemical shift perturbation can also be used to localize the binding site on the protein surface.

The fairly diluted samples required (0.2-0.3mM) and the short acquisition time required are two of the main advantages of this experiment.

2.4.3 ^{15}N, ^{13}C rejected NOESY

^{15}N and ^{13}C isotope filtering is a very useful approach for obtaining three-dimensional information about protein-protein and protein-ligand structures. (Zwahlen C., et al. 1997) ^{15}N, ^{13}C rejected NOESY experiments are usually run on selectively isotope labelled complexes where just one component of the complex is labelled. These experiments are NOE based experiments in which the cross-peaks arise exclusively from intra-molecular

proton-proton NOE interactions. All isotope filters rely on the presence or the absence of a scalar coupling between a proton and a hetero-nucleus X. If for example the X nucleus is nitrogen the coupling to the heteronucleus can be used to achieve discrimination between the two different classes of protons (^{1}H-^{15}N and ^{1}H-^{14}N). In fact it is possible to generate heteronuclear antiphase spin coherence in one case but not in the other. The antiphase coherence can then be manipulated so that it is either preserved or eliminated and this allows one to discriminate between the two classes of protons.^{15}N filtering, because of the relative uniformity of the $^{1}J_{HN}$ coupling constant (~ 90 Hz for the backbone amide groups) is less technically demanding than ^{13}C filtering where the $^{1}J_{CH}$ coupling constants vary much more (from ~ 125 Hz for the aliphatic methyl group to ~220 Hz for the imidazole ring in histidine). Is also worth mentioning that isotope filtered experiment are intrinsically less sensitive than non-filtered ones because of the increased pulse length associated with the incorporation of filter delays.

CHAPTER 3

RESULTS: the structure of holo F1

3. RESULTS: the structure of holo F1

3.1 Comparison of TnC sequences

As a prerequisite to the structural studies we first started by comparing the amino-acid sequence and the calcium binding feature of F1 to the other known TnCs.

The sequences of TnCs are well conserved across vertebrates and invertebrates; the calcium binding capacity varies more significantly among the many isoforms. For example, skeletal TnC can bind four calcium ions, one in each EF-hand motif, (Potter *et al.*,1977), cardiac TnC can bind three calcium ions: one of the EF hand loops in the N terminal domain does not bind calcium. (Van Eerd *et al.*,1975) In insect IFM two TnC isoforms are present: F1 and F2: F1 has only one EF hand motif active in the C-terminal domain, F2 on the other hand has two, one in each domain.(Qiu *et al.*,2003)

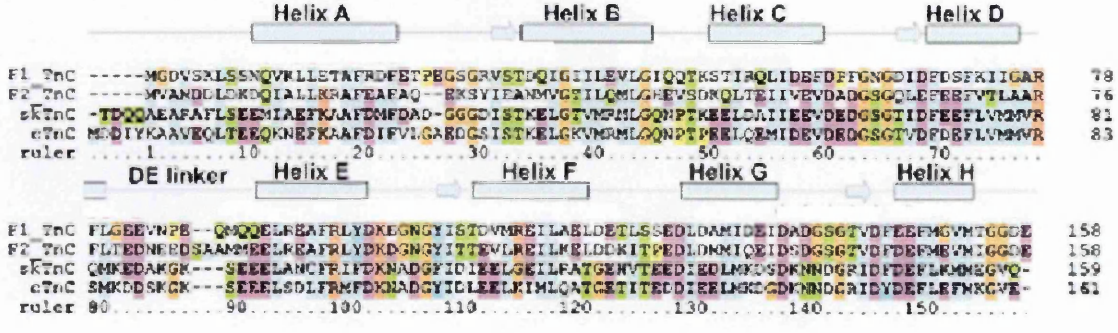


Fig. 3.1.1 Sequence alignment and the secondary structure elements of F1,F2, skeletal and cardiac TnC isoforms. The squares represent α-helical structures, the arrows β strand structures.

3.2 Description of F1 production

The recombinant protein was expressed in *E. coli.* (see chapter 8 for details) The protein is over expressed with yields of some 4-8mg per litre of culture depending on what media are used to grow the cells. The protein is expressed in the soluble fraction and can be concentrated up to 1mM in 100mM salt without any precipitation occurring. Affinity and size exclusion chromatography were used to purify it (see chapter 8.1).

3.3 Description of F1 thermal stability

The thermal stabilities of the apo and holo forms of F1 TnC were tested. Thermal unfolding curves, as recorded by circular dichroism (CD) on the apo form, show that the protein undergoes a highly cooperative transition with an apparent melting point around 50°C. If however the curve is fit with a single component model, a poor chi-squared value is found (2.34), suggesting the presence of a two-step unfolding process. When a two component fit was used the chi-squared value fell to 1.03. The melting temperatures for the two component fits were 37.5°C and 62.4°C.

Holo F1 TnC undergoes what is apparently a single transition at a temperature of 56.5°C, indicating that calcium stabilizes the C-terminal domain and leads to an unfolding temperature comparable to that of the N-terminal domain. These results indicate that, although the apo C-terminal domain is folded at room temperature, it has a significantly lower stability than the N-terminal domain. The difference is dramatically reduced in the holo form.

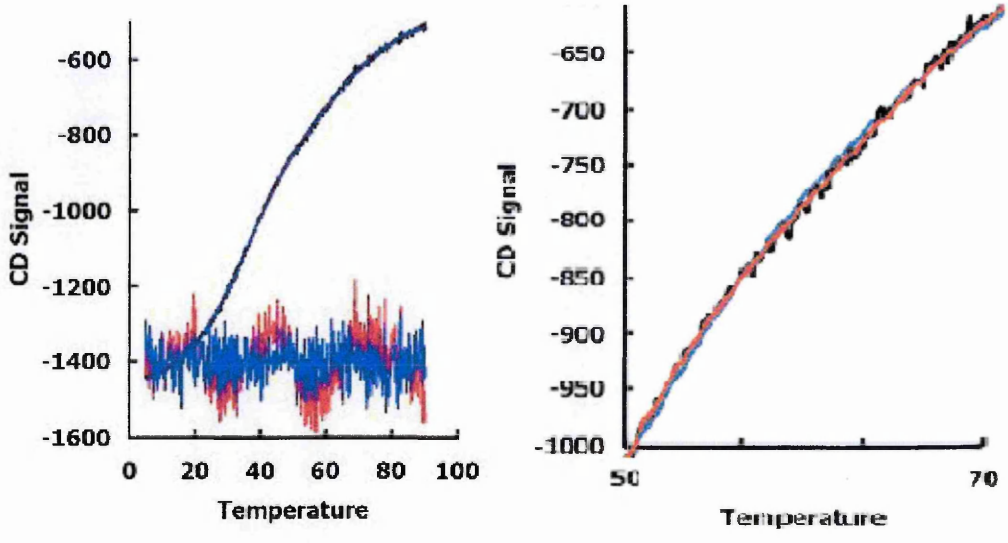


Fig. 3.3.1. Thermal unfolding of apo F1 monitored at 222nm. The left hand plot shows the full curve with a one component fit (in blue) and a two component fit (in red). This plot also shows that the residuals (multiplied by 10 for clarity) are highly non-random for the one component fit. The poor quality of the one component fit is also reflected in the chi-squared values (1.03 for the two component fit and 2.34 for the one component fit). The right hand plot just shows an expanded view. The Tm values for the two component fit are 37.5 °C and 62.4. °C

3.4 Description of the holo F1 structure calculation

F1 gives a good well dispersed {^{1}H,^{15}N} HSQC where all but one of the non proline signals are identifiable. The experiments for the backbone and side chain assignments and the NOESY based experiments for the structure calculation worked reasonably well. The 3D ^{15}N-edited TOCSY was the only exception. (Fesik and Zuiderweg, 1988) Assigning the spin systems of the residues belonging to the flexible linker was particularly time consuming because of the severe overlap. The restraints obtained through the NOESY based experiments were enough for the structure to converge during the structure calculation and to produce structures with a reasonably low root mean square deviation (r.m.s.d.) ~0.9 Å. The 98% complete assignment is accessible at the BMRB webpage (accession code: 6081). http://www.bmrb.wisc.edu/data_library/generate_summary.php?bmrbId=6081

3.5 Description of the holo F1 TnC structure

The solution structure of holo F1 TnC consists of a typical EF hand fold with two globular domains spaced by a flexible linker ten residues long (figure 3.5.1). The two domains superimpose independently with a backbone r.m.s.d. of ~ 0.9 Å. This low value reflects the fact that each of the two domains has a compact and well defined fold lacking long unstructured termini or loops. Each domain consists of two EF hand motifs, each of which comprises two α-helices flanking the twelve residue loop. The two loops form the typical short antiparallel β-sheet which keeps the two EF-hands side by side. The N helix which precedes the first helix in the N domain is short and is formed only in some of the structures.

In the solution structure of *Lethocerus* F1 TnC the linker is unstructured and highly flexible allowing the domains to tumble independently, as also shown by the relaxation data. (see chapter 3.7)

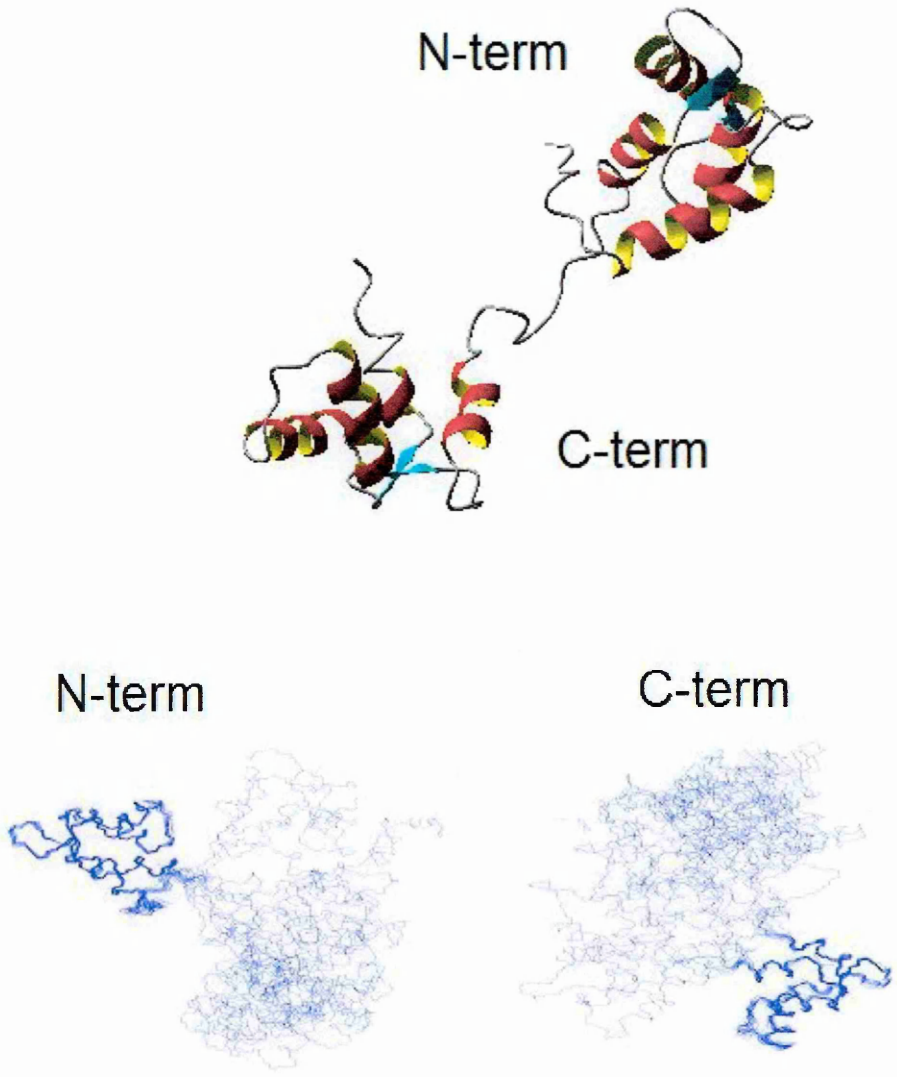


Fig. 3.5.1 Ribbon representation of the solution structure of holo F1 and the three dimensional fold of F1 represented by the NMR bundle. The 10 lowest energy structures of F1 resulting from Aria calculations, after energy minimization in a shell of water, are displayed.

3.6 Quality checks on the structures

The structure calculation was done using the ARIA program.(Linge *et al.*,2003) Procheck (Laskowski *et al.*, 1996) and Whatif (Vriend, 1990) were used to assess the quality of the structures calculated.

Final NMR Restraints	
Total distance restraints	3840
Unambiguous/ambiguous	3401/439
Intraresidue	1578
Sequential	840
Medium (residue i to i + j, j = 2 – 4)	573
Long-range (residue i to I + j, j > 4)	849
Deviation from Idealized Geometry	
Bond lengths (Å)	0.003 ± 0.000
Bond angles (°)	0.396 ± 0.023
Improper dihedrals (°)	0.296 ± 0.023
Restraint Violations	
Distance restraint violation > 0.5 Å	0
Dihedral restraint violation > 5°	0
Coordinate Precision (Å) with Respect to the Mean Structure	
Backbone of N-terminal structured regions	0.35
Heavy atoms of N-terminal structured regions	0.8
Backbone of C-terminal structured regions	0.61
Heavy atoms of C-terminal structured regions	0.95
Backbone of N-terminal secondary structure element	0.26
Heavy atoms of N-terminal secondary structure element	0.79
Backbone of C-terminal secondary structure elements	0.34
Heavy atoms of C-terminal secondary structure element	0.82
Whatif Quality Check	
First generation packing quality	- 1.648
Second generation packing quality	-1.058
Ramachandran plot appearance	-2.960
$\chi1$-$\chi2$ rotamer normality	-2.205
Backbone conformation	-1.915
Procheck Ramachandran Statistics	
Most favored region (%)	83.6
Additionally allowed regions (%)	14.4
Generously allowed regions (%)	0.7
Disallowed regions (%)	1.2

Table 3.6.1 The statistics calculated for the 10 lowest energy structures of holo F1 after water refinement are reported

3.7 Description of holo F1 dynamics

In the solution structure of *Lethocerus* holo F1 TnC the linker between the two domains is unstructured and highly flexible, allowing the domains to tumble independently. This result is confirmed by the analysis of the dynamics data and by the absence of NOE signals between the two domains. It is also worth noting that the C-terminal domain exhibits a slightly higher flexibility when compared to the N terminal domain: on average its T_1 values are longer and T_2 ones shorter. The analysis of T_1,T_2 and NOE is shown in fig 3.7.1

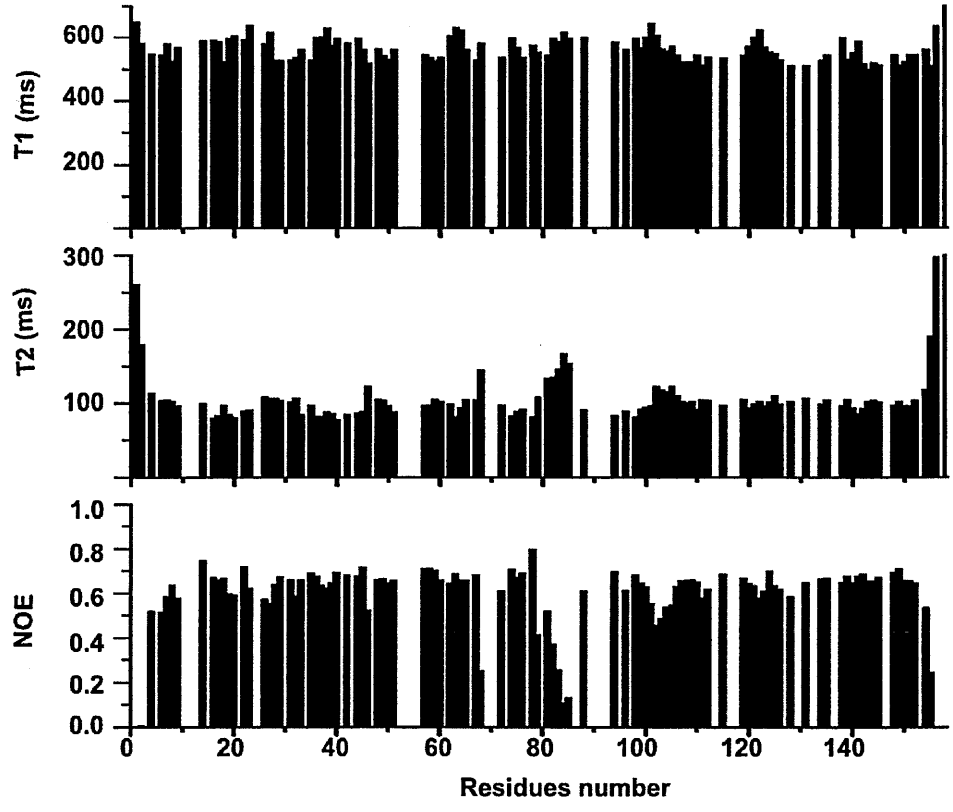


Fig. 3.7.1 Relaxation parameters of the backbone ^{15}N nuclei plotted as a function of the residue number. The data were recorded at 25°C and 500 MHz.

3.8 F1 calcium binding features

Lethocerus F1 is able to bind just one calcium ion in the second EF hand motif in the C-terminal domain. The stoichiometry previously obtained by mass spectrometry (Qiu et al., 2003) and the identity of the calcium binding site on F1 TnC were confirmed by NMR. This technique is able to provide a direct identification of the calcium binding sites in EF-hands due to the presence of a highly diagnostic value for the ^{15}N chemical shift of the amido nitrogen of the residue in position 8 of a canonical EF-hand loop (Biekofsky *et al.*, 1998). In F1 this position corresponds to residue Val 144 and comparison of the {^{1}H,^{15}N} HSQC spectra of the apo and holo forms of F1 TnC shows that V144 moves from 8.99 and 120.0 ppm to 9.89 and 128.8 ppm. This observation proves conclusively that a single calcium ion is accommodated in site IV of the C-terminal domain, whereas the other three coordination sites are inactive even at high (millimolar) concentrations of this cation. The calcium affinity constants for F1 were measured using 5,5'-Br_2BAPTA at 25 °C and 40 °C and the K_d values were 0.62 ± 0.06 μM and 1.70 ± 0.28 μM respectively.

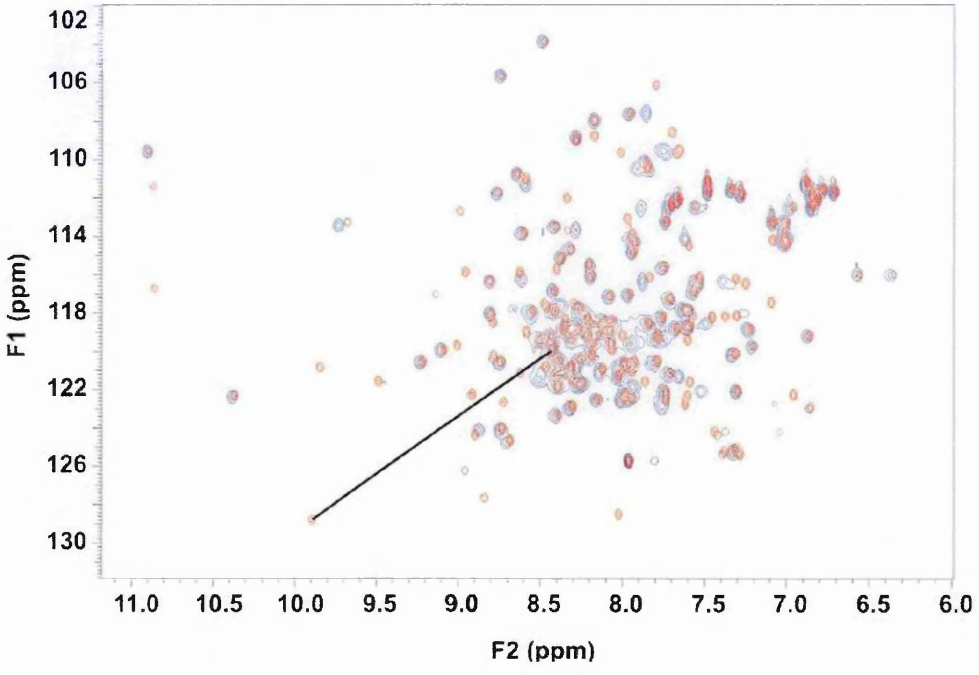


Fig. 3.8.1 Comparison of the HSQC spectra of apo (blue) and holo (red) *Lethocerus* F1 TnC. The spectra were recorded at 25°C on a 600 MHz spectrometer. The shift upon Ca^{2+} addition of the resonance of V144 is indicated.

3.9 Structural comparison with other TnCs

In general two EF hand motifs can have an "open" or "closed" conformation based upon the inter-helical angles that the helices flanking the calcium binding loop make with respect to each other. In F1 the N-terminal domain does not bind calcium and the overall arrangement of the four helices is such as to have a closed conformation, as is the case with most of the known EF-hands in the apo-form.

The F1 C-terminal domain is the only C-terminal TnC structure so far solved which binds a single calcium in its C-terminal domain. In the presence of one calcium ion, the domain has an open conformation. This is at variance with what happens in the N terminal domain of cardiac TnC. The N domain of cardiac TnC is made of two EF-hand motifs and only one of them binds calcium as in F1, but upon binding the domain does not undergo significant structural changes to produce an "open" conformation. This domain is in a closed conformation regardless of the presence of calcium. The presence of both: calcium and the TnI regulatory region appears to be necessary to stabilize a more "open" conformation with increase in the exposure of the hydrophobic pocket. In this case binding of calcium causes only an increase in the affinity for TnI compared to the apo form but not a significant change in the structure of the unbound form. (Li *et al.*, 1999, Dong *et al.*, 1999)

In holo F1 the C-terminal domain has an open conformation. Submitting the F1 C-terminal domain structure to the Dali server against the PDB database it picks up the structure of C-domain of skeletal TnC in complex with the correspondent region of TnI as the most similar. It seems as if the F1 C-terminal domain, even without the peptide, is already arranged to interact with the N-terminus region of TnI.

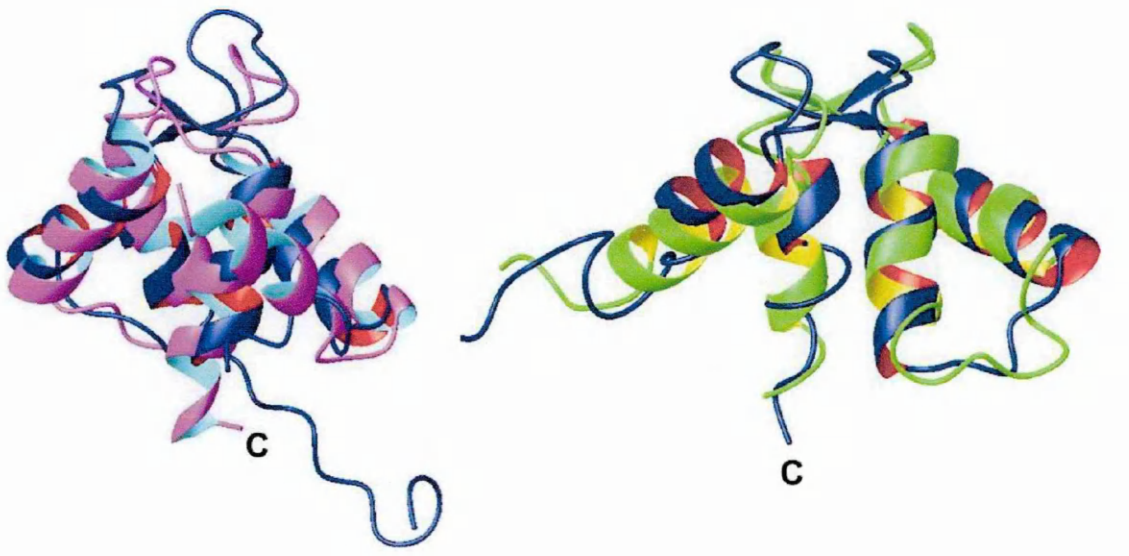


Fig. 3.9.1 Conformation of the two domains of holo F1 TnC. Structural comparison of the two domains of F1 TnC with the PDB picks up as the highest hits for the N terminal domain the apo N-terminal domain of skeletal TnC, which is in closed conformation (PDB entry: 1skt, Tsuda *et al.,* 1999) and the holo C-terminal domain of cardiac TnC in complex with the corresponding N-terminus of TnI for the C terminal domain (1ozs, Lindhout and Sykes, 2003).

CHAPTER 4

RESULTS: the structure of apo F1

4. RESULTS: the structure of apo F1

4.1 Why solve the structure of apo F1?

The finding that the holo F1 C terminal domain in the presence of calcium is in an open conformation resembles the behaviour of skeletal TnC and calmodulin (Slupsky and Sykes, 1995, Zhang *et al.,* 1995, Mercier *et al.,* 2001) but it is at variance with what is observed in the cardiac TnC N-terminal domain. (Li *et al.,*1999; Dong *et al.,*1999). In order to elucidate the role of calcium in the three-dimensional structure of F1 the solution structure of F1 in the apo form was solved. The aim was to find out if calcium causes the F1 C-terminal domain to be in an open conformation, such as in skeletal TnC and calmodulin, or if the domain is itself in an open conformation regardless of the presence of calcium, therefore being, from a structural point of view, calcium insensitive.

Considering that in the F1 holo structure the two domains tumble independently, do not have any contact, and that the N-terminal domain does not bind calcium, it is reasonable to assume that in the absence of calcium the N-terminal domain has the same structure as in the presence of calcium. In order to speed up and simplify the task of solving the structure of the apo form of F1 the C domain was subcloned and only the apo C terminal domain structure was solved.

4.2 Description of the protein production

The F1 C terminal domain of F1 was sub cloned in a pET24d (M11) expression vector (Novagen), the protein was expressed in *E. coli.* (see chapter 8 for details) The protein is generally over expressed at the level of 4-8mg per litre of culture, depending on what media are used to grow the cells. The protein is expressed in the soluble fraction and can be concentrated up to 1mM in 100mM salt without any precipitation occurring. Affinity and size exclusion chromatography were used to purify it.(see chapter 8.2)

4.3 Description of F1 C-terminal domain stability

The thermal stability of the apo and holo forms of C-terminal domains of F1 were tested. Thermal unfolding curves, as recorded by CD on the two form, show in both cases a

highly cooperative transition with a melting point around 49°C, for the apo form and around 76° C for the holo form.

4.4 Description of F1 C-terminal domain secondary structure content

CD spectra of the apo and holo C domain were acquired. The curves were de-convoluted and the secondary structure content was calculated. The holo and the apo forms have very similar secondary structure contents.

	α-helix	β-sheet	turn	Random
apo C-F1	50.0%	7.0%	15.5%	27.5%
holo C-F1	53.5%	7.5%	13.0%	26.0%

Tab. 4.4.1 The secondary structure content for apo and holo F1 C-terminal domains.

4.5 Description of the apo F1 structure calculation

The F1 C domain does not give a good {^{1}H,^{15}N} HSQC spectrum at 25°C and at pH 6.8. Only 65 of the 73 residues are present. Different pH and temperature conditions were tried to improve the quality of the spectrum but there was not a significant improvement. The usual experiments for the backbone and side chain assignments and for the structure calculation were acquired. The ^{15}N-edited TOCSY and the ^{15}N edited NOESY-HSQC (Fesik and Zuiderweg, 1988) were of particularly poor quality. The restraints obtained through the NOESY based experiments turned out to be enough for the structure to converge during the structure calculation and to produce structures with a backbone r.m.s.d. of ~1.5 Å.

4.6 Apo and holo C terminus F1 ^{1}H-^{15}N HSQC spectra comparison

The F1 C terminal domain {^{1}H,^{15}N}HSQC spectra recorded in the presence and in the absence of calcium were compared. Besides the chemical shift changes of residues directly involved in the binding of calcium, the two spectra do not look significantly different and are both typical of folded species, suggesting that the domain is also intrinsically structured in the absence of calcium. Moreover the apo and holo C-terminal

domain HSQC peaks overlap almost perfectly to the corresponding residues in the full length apo and holo F1 respectively. This observation confirms indirectly that in full length F1 the two domains tumble independently in solution

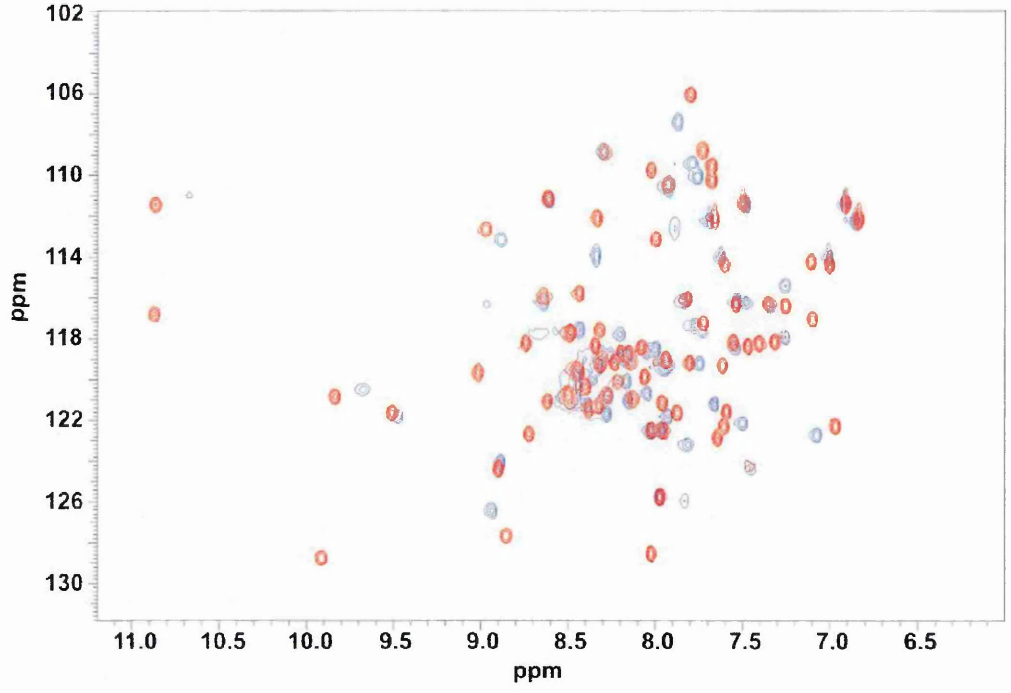


Fig. 4.6.1 Superposition of the HSQC spectra of apo (blue) and holo (red) forms of the C-terminal domains of F1 TnC. The spectra are both those of well folded proteins, although the apo form is presumably less rigid than the holo form, as suggested by the broadening of some resonances and by the absence of others. The two spectra superpose almost perfectly to those of the full-length protein indicating that the two domains do not have significant interactions with each others.

4.7 Description of apo F1 C-terminal domain structure

The solution structure of apo F1 C domain consists of a typical EF hand fold. The domain consists of two EF hand motifs, each of which comprises two α-helices flanking the eleven residues loop. The two loops form the typical short antiparallel β-sheet which keeps the two EF-hands motifs side by side. The backbone r.m.s.d. for the ten lowest energy structures is 1.5 Å. The value reflects the lack of restraints, especially for the residues belonging to the loops connecting the four helices.

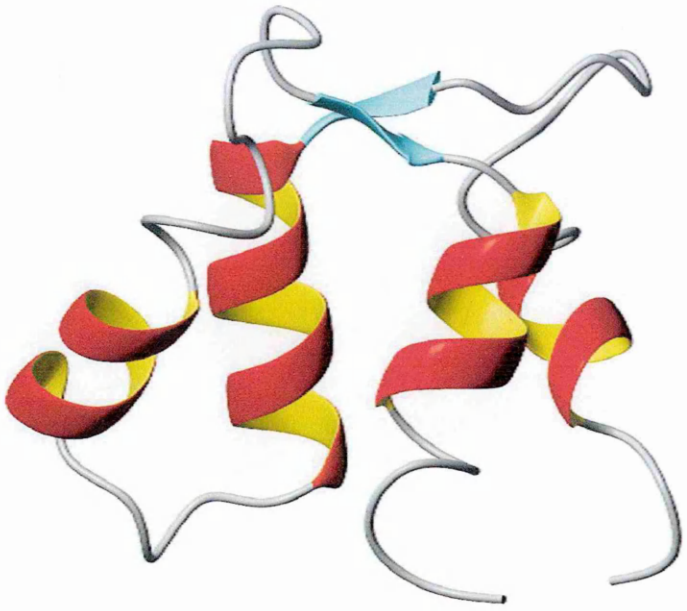


Fig. 4.7.1 Ribbon representation of the solution structure of apo F1 C-terminal domain

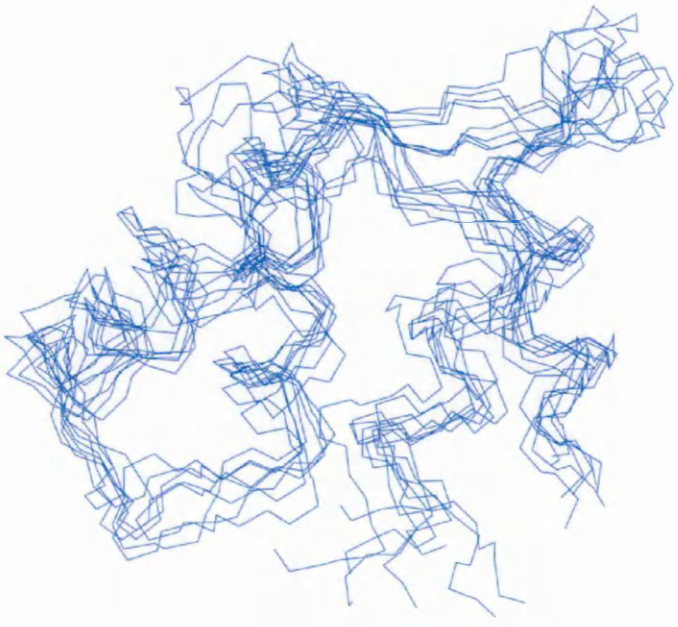


Fig. 4.7.2 Three dimensional fold of F1 C-terminal domain represented by the NMR bundle. The 10 lowest energy structures of F1 resulting from ARIA calculations, after energy minimization in a shell of water, are displayed

4.8 Quality check of the structure

The structure calculation was performed using ARIA program.(Linge *et al.*,2003) Procheck (Laskowski *et al.*, 1996) and Whatif (Vriend, 1990) were used to asses the quality of the structures calculated.

Final NMR Restraints	
Total distance restraints	1230
Unambiguous/ambiguous	1023/207
Intraresidue	609
Sequential	249
Medium (residue i to i + j, j = 2 – 4)	132
Long-range (residue i to I + j, j > 4)	240
Deviation from Idealized Geometry	
Bond lengths (Å)	0.004±0.000
Bond angles (°)	0.452±0.013
Improper dihedrals (°)	0.316±0.123
Restraint Violations	
Distance restraint violation > 0.5 Å	0
Dihedral restraint violation > 5°	0
Coordinate Precision (Å) with Respect to the Mean Structure	
Backbone of apo F1C-terminal structured regions	1.50
Heavy atoms of apo F1 C-terminal structured regions	1.97
Backbone of apo F1 C-terminal secondary structure elements	0.97
Heavy atoms of apo F1 C-terminal secondary structure element	1.49
Whatif Quality Check	
First generation packing quality	-2.27
Second generation packing quality	-2.39
Ramachandran plot appearance	-3.46
$\chi1$-$\chi2$ rotamer normality	-0.60
Backbone conformation	-1.72
Procheck Ramachandran Statistics	
Most favored region (%)	82.0
Additionally allowed regions (%)	14.8
Generously allowed regions (%)	1.6
Disallowed regions (%)	1.6

Table 4.8.1 The statistics calculated for the 10 lowest energy structures of apo F1 C terminal domain after water refinement are reported.

4.9 Structural comparison between the apo and holo structures

When the apo and holo structures of F1 C-terminal domain are superimposed three main features are evident. First there is no significant difference in the secondary structure elements, confirming the CD data, second the angles between the helical axes are roughly the same and finally the loops connecting the helices are much more flexible in the apo form than in the holo form. Calcium binding in this system seems to cause a decrease in flexibility in the loops but not any major three dimensional rearrangement.

The F1 C-terminal domain is therefore structurally calcium insensitive. This feature is similar to the N-terminal domain of cardiac TnC: in both cases calcium does not cause any major structural change. In cardiac TnC the domain stays closed in the presence and in the absence of calcium whereas in F1 it stays open regardless of the presence of calcium. (Li *et al.*, 1999, Dong *et al.*, 1999)

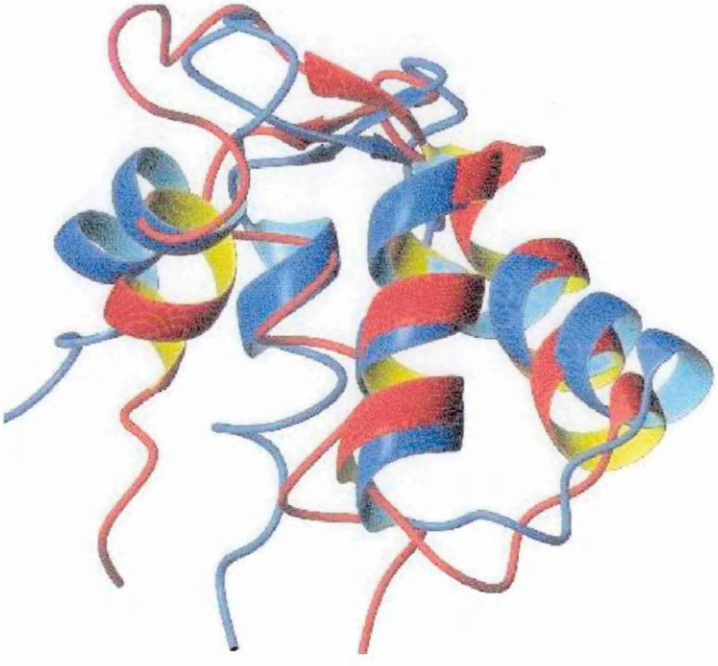


Fig. 4.9.1 Superimposition of the apo (red) and holo (blue) F1 C-terminal domains. The secondary structure elements and the overall three dimensional arrangement between the two structure are very similar.

Chapter 5

RESULTS: Result: F1 and TnH(30-61), TnH(126-159) peptides interaction

5. Results: Interactions between F1 and TnH(30-61), TnH(126-159) peptides

5.1 TnC and TnH interaction

We used the structural information gained on F1 to characterize the interactions of holo and apo F1 with TnH. We first tried to identify regions of the *Lethocerus* TnH sequence which should correspond to the N terminal region and the switch peptide in vertebrate muscles and are therefore expected to interact with the C and N-terminal domains of TnC respectively.

TnH is more complex than vertebrate TnI. A region which shares some homology (around 27%) with vertebrate TnI is sandwiched between an N-terminal region and a proline-and-alanine-rich extension, which is insect specific. The N-terminal region (30-61) and the inhibitory peptide share sufficient sequence similarity to suggest that their role may also be preserved in *Lethocerus*, whereas the switch peptide is shorter and highly divergent. Arg115 in TnI, for instance, which in skeletal muscle is salt bridged with TnC Glu63 and Gln84 is substituted by a serine (Ser147). Likewise there is no conservation of the residues of TnC which should contribute to the interaction.

These observations suggest a highly divergent regulation of contraction in stretch regulated insect muscles.

To test this hypothesis, we designed two synthetic peptides, TnH(30-61) and TnH(126-159), which were used in binding studies with the apo and holo forms of F1 TnC.

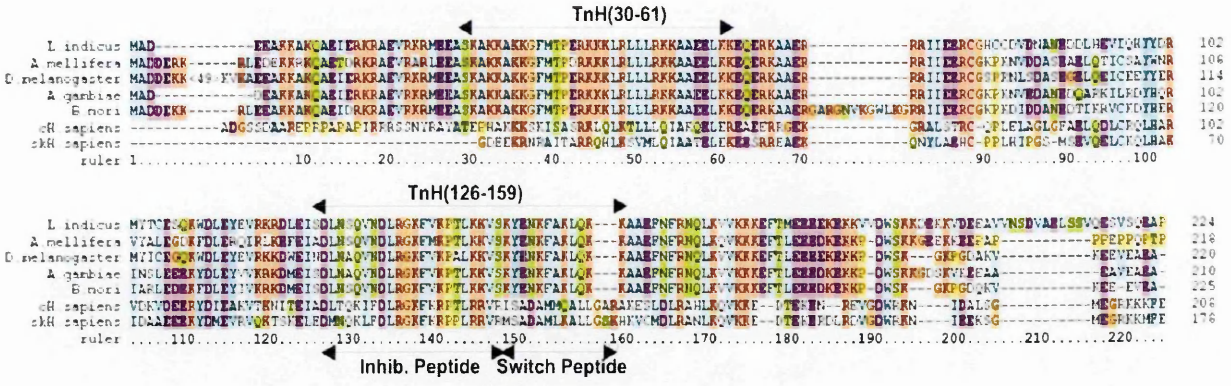


Fig. 5.1.1 Insect TnI sequences are compared with those from vertebrates. The positions of the peptides used in this study are indicated. The region corresponding to the switch peptide is highly divergent in insects compared to vertebrates. The *Lethocerus* TnH sequence is truncated at the end of the region homologous to vertebrate TnI; the full length sequence is 484 residues. The GenBank/EMBL accession number of *Lethocerus* TnH is AJ621044.

5.2 Fluorescence titrations

F1, TnH(30-61) and TnH(126-129) are spectroscopically silent (i.e., they do not contain tryptophans). It is therefore not possible to measure F1 TnH peptides affinity constants through direct fluorometric titrations. In order to circumvent this limitation competition assays were used to measure these constants. This method is based on the use of a tryptophan containing peptide, in this case melittin, which competes for the same binding site of TnH(30-61) or TnH(126-159) on F1 surface. Melittin is a 26 amino acid long peptide containing a tryptophan in position 19. Melittin is known to bind to F1 and EF hand protein in general with a 1:1 stoichiometry (Comte, M *et al.*, 1983)

5.2.1 Direct fluorescence titration

In order to determine F1 TnH(30-61) and F1 TnH(126-159) affinity constants through competition assays the first step was to determine through direct fluorescence titration the binding affinity of melittin for F1 in the apo and holo forms.

5.2.2 Apo and holo F1 melittin titration

The experiments were done by titrating a solution of melittin (usually 1-5 μM) with apo F1 or holo F1 stock solutions. The fluorescence intensity at 322 nm was recorded and a control titration of F1 without melittin was performed in order to obtain an accurate value of the F1 contribution to the fluorescence deriving from the two tyrosines present in its sequence. The experiments were performed at 25°C and 40°C and gave a result of 26±7nM for the K_d of apo F1 and melittin at 25°C and a K_d of 95±18nM for apo F1 and melittin at 40°C. The values of the K_d for holo F1 and melittin at 25°C and 40°C were 5.2±1.2nM and 12.5±4nM respectively.

5.2.3 Apo and holo F1 TnH(30-61) competition assay

First a melittin solution was titrated in with a stock solution of F1 and the fluorescence intensity was recorded at 322nm. The concentrations were chosen such as to have [melittin] > [F1] in order to make sure that there was little F1 free during the titration, then once a plateau was reached small aliquots of the concentrated TnH(30-61) (usually 30-50μM) were added to the solution and the fluorescence at 322nm was measured. A control titration without melittin was done to determine the specific fluorescence of F1 and of the F1/TnH(30-61) complex. The same experiment were done at 25°C and 40°C, using apo and holo F1 and gave a result of 1.3±0.3nM for the K_d of apo F1 and TnH(30-61) at 25°C and a K_d of 3.9±1.2nM for apo F1 and TnH(30-61) at 40°C. The values of the K_d for holo F1 and TnH(30-61) at 25°C and 40°C were 1.5±0.5nM and 2.8±1.1nM respectively

5.2.4 Apo and holo F1 TnH(126-159) competition assay

The experiments were done by titrating melittin with a stock solution of F1 in the presence of varying excesses of TnH(126-159).The fluorescence intensity was measured at 322nm an whenever possible the titrations were continued until a constant fluorescence value was reached. A control titration without melittin was done to determine the specific fluorescence of F1 and of the F1/TnH(126-159) complex. The same experiments were done at 25°C and 40°C using apo and holo F1 and gave a result of 5.2±1.4μM for the K_d of apo F1 and TnH(126-159) at 25°C and a K_d of 42±9μM for apo F1 and TnH(126-159)

at 40°C. The values of the K_d for holo F1 and TnH(126-159) at 25°C and 40°C were 8.6±3μM and 18±6μM respectively.

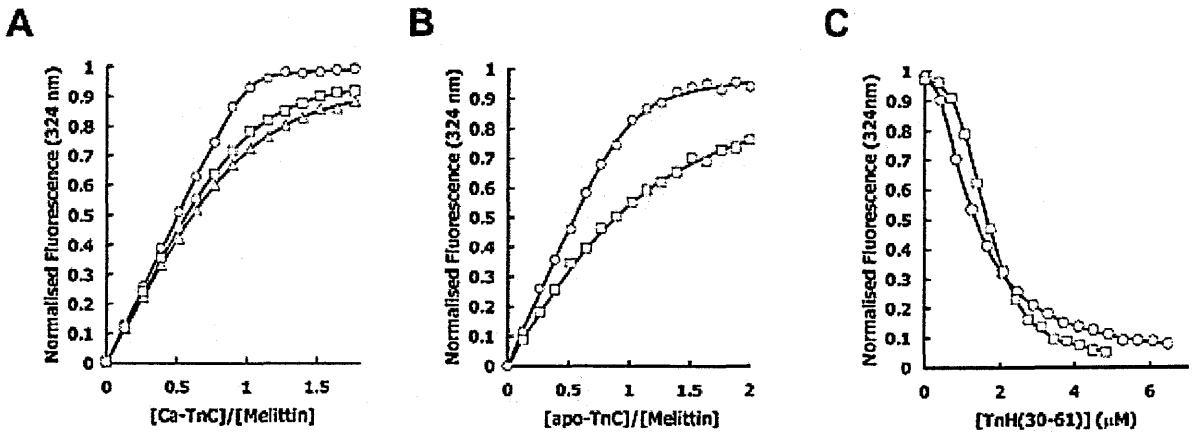


Figure 5.2.4.1 Probing the interactions of synthetic peptides spanning the TnH sequence. A) Titration of melittin (0.65 μM) with holo F1 TnC in the absence (circles) and presence of 37.5 (squares) or 75 uM (triangles) TnH(126-159). B) Titration of melittin (0.65 μM) with apo F1 TnC in the absence (circles) and in the presence (squares) of 75 μM TnH(126-159). C) Titration of 1.6 μM melitin plus 1.9 μM holo F1 TnC (circles) or 2.4 μM apo TnC (squares) with TnH(30-61).

5.3 Apo and holo F1 TnH(126-159) far-UV CD spectra

The far-UV CD spectra of the complexes of apo and holo F1 with TnH(126-159) are significantly more intense (higher α-helical content) than the sum of the appropriate component spectra. We thus performed titrations of F1 with TnH(126-159) in which we monitored the far-UV CD signal change accompanying complex formation. Although the relatively high concentrations required for these experiments did not permit determination of precise K_d values, they clearly showed that the interaction is similarly strong, both in the presence and in the absence of calcium.

5.4 Apo and holo F1 TnH(126-159) Chemical and Thermal unfolding

To have an independent measurement of the K_d for TnH(126-159), the TnH peptide with lower affinity, we studied its effects on the stability of the C-terminal domain of apo F1 to chemical denaturation. We thus carried out a chemical denaturation experiment on the apo and holo F1 in the absence and in the presence of an excess of TnH(126-159). Analysis of the urea unfolding data gave a stability increase consistent with dissociation constants in the range 10-15 μM.

We also checked the effect of an excess of TnH(126-159) on thermal denaturation for the apo and holo F1. The mid-point for the unfolding of the C-terminal domain of apo F1 increased from 37.5 to 52.8°C in the presence of an excess of TnH(126-159), whilst the increase was from 56.5 to 62.7°C for holo TnC. These increases indicate that the interaction is in both cases mostly of hydrophobic nature.

5.5 F1 Calcium affinities in the presence and in the absence of TnH peptides

To check if calcium has a regulatory role in the peptide binding affinities, the F1 calcium affinities in the absence and in the presence of either TnH peptides were measured using 5,5'-Br_2-BAPTA: a calcium chelator which changes its extinction coefficient when passing from the free to the calcium bound form.

The basic idea behind this experiments is that if any of the peptides binds more strongly in the presence of calcium, then calcium should have a higher affinity for the apo protein in the presence of that peptide.

In fact according to the scheme below, considering that both reaction pathways give the same product, the ratio K_1K_2 / K_3K_4 has to remain constant.

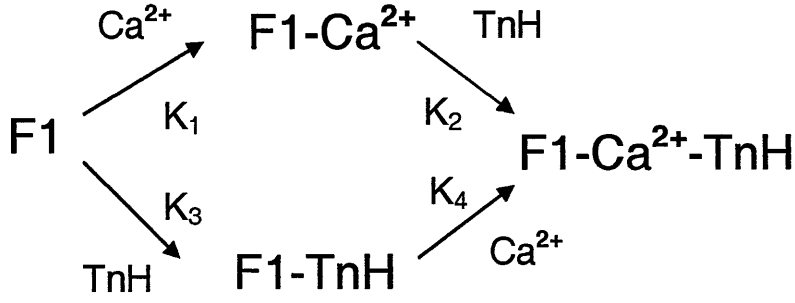


Fig. 5.5.1 The F1,Ca^{2+} and TnH binding cycle. The fact that the K_1K_2 / K_3K_4 ratio has to remain constant allows an indirect way to check the possible regulatory role of Ca^{2+} TnH in the scheme represent either TnH(30-61) or TnH(126-159).

The experiments were done by titrating a mixture of chelator and apo F1 with a stock solution of calcium. The absorption at 263nm was recorded and the data were acquired at 25°C and 40°C. The aim of the experiment was to measure K_1 and K_4 as defined in the figure 5.5.1. Calcium K_d (K_1) was 0.62 ± 0.06 μM at 25°C and 1.70 ± 0.28 μM at 40°C. The calcium K_d in the presence of an excess of TnH(30-61) (K_4) was 0.66 ± 0.08 μM at 25°C and 1.19 ± 0.13 μM at 40°C, whereas the calcium K_d (K_4) in presence of an excess of TnH(126-159) (K_4) was 0.42 ± 0.05 μM at 25°C and 0.71 ± 0.09 μM at 40°C.

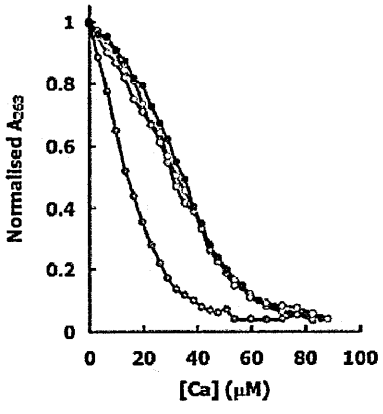


Fig 5.5.2 Probing Ca^{2+} affinity towards F1 in the presence and in the absence of TnH(30-61) and TnH(126-159). Titration of 5,5-Br_2BAPTA with Ca^{2+}(red circles) in the presence of 33.5 μM of F1, (green circles) and of 200 μM of TnH(30-61) (blue circles) or 200 μM of TnH(126-159) (cyan circles).

5.6 F1 TnH(30-60) NMR titration

A ^{15}N F1 solution was titrated with a stock solution of TnH(30-61). The titration of ^{15}N F1 with TnI(30-61) was monitored by two dimensional {^{1}H,^{15}N}-HSQC-NMR. The experiments were done using both apo and holo F1. In both cases the chemical shift variations fall into the slow exchange limit in the NMR time scale where the bound and unbound species give rise to a cross peak each and their intensity is proportional to the concentration of the two species in solution. In this case the intermediate species are not observed so it is possible to see only the initial and final step of the titration. The peaks perturbed belong to the C-terminal domain of F1 and they are the same regardless of the presence of calcium.

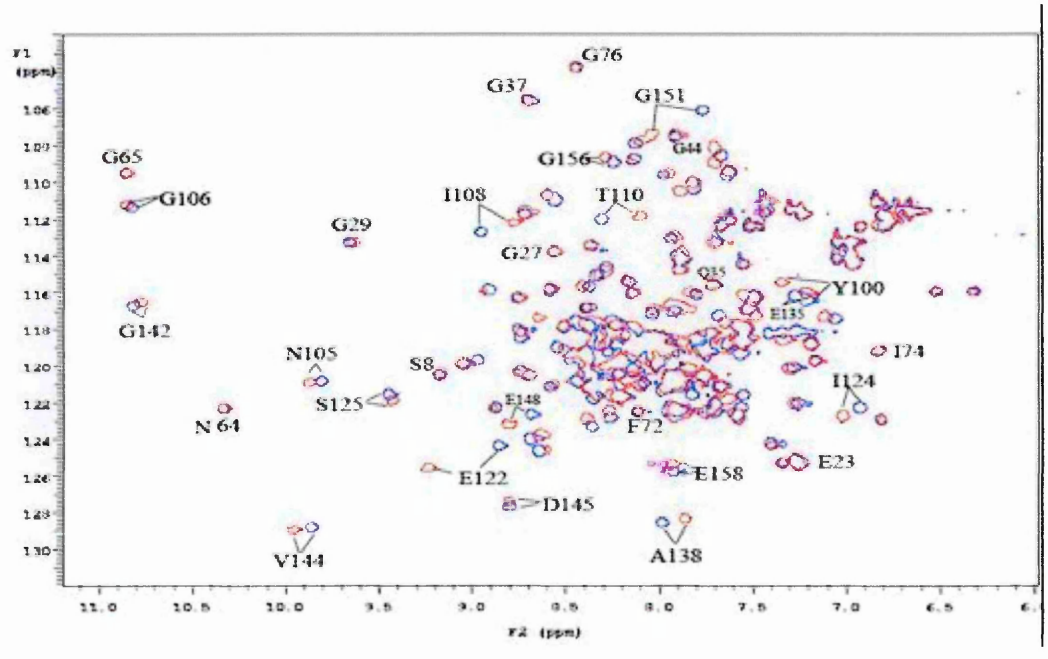


Fig. 5.6.1 Superposition of the HSQC spectra of ^{15}N labelled holo F1 in the absence (red) and in the presence of a molar excess (~1:3) of peptides (blue). The spectra were recorded at 25°C and 600 MHz. Some of the residues perturbed are labelled.

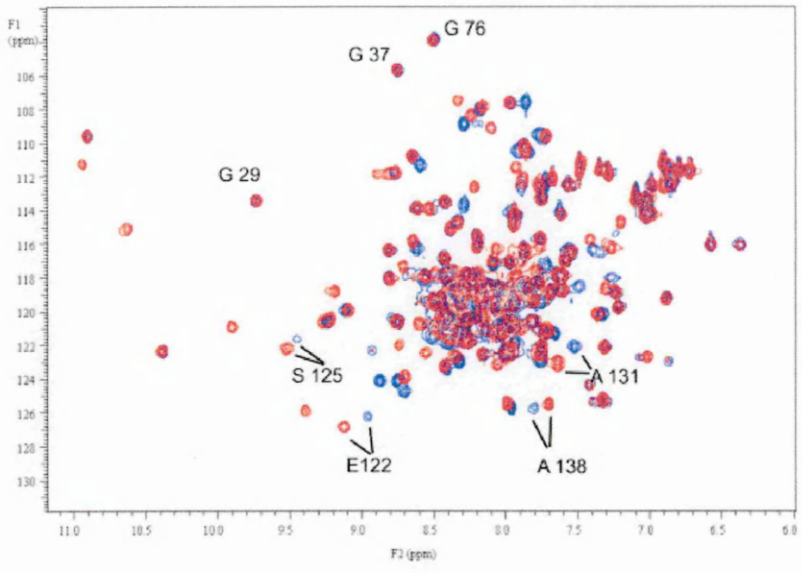


Fig. 5.6.2 Superposition of the HSQC spectra of ^{15}N labelled apo F1 in the absence (blue) and in the presence of a molar excess (~1:3) of peptides (red). The spectra were recorded at 25°C and 600 MHz. Some of the residues perturbed are labelled.

5.7 Holo F1 TnH(126-159) NMR titration

A ^{15}N F1 solution was titrated with a stock solution of TnH(126-159). The titration of ^{15}N F1 with TnH(126-159) was monitored by two dimensional {^{1}H,^{15}N}-HSQC-NMR. The chemical shift changes fall into the fast exchange limit in the NMR timescale where it is possible to observe at each point of the titration a cross peak which is the weighted average of the bound and unbound species. The cross peaks are the intermediate species and in this case is possible to follow the chemical shifts perturbation at every single point of the titration. In this case only holo F1 was used during the titration. The peaks perturbed all belong to the C-terminal domain of F1 and most of them are the same as the ones in the TnH(30-61) titration. These data seem to suggest that on F1 C-terminal domain surface there is a single TnH binding site that is able to accommodate either peptides.

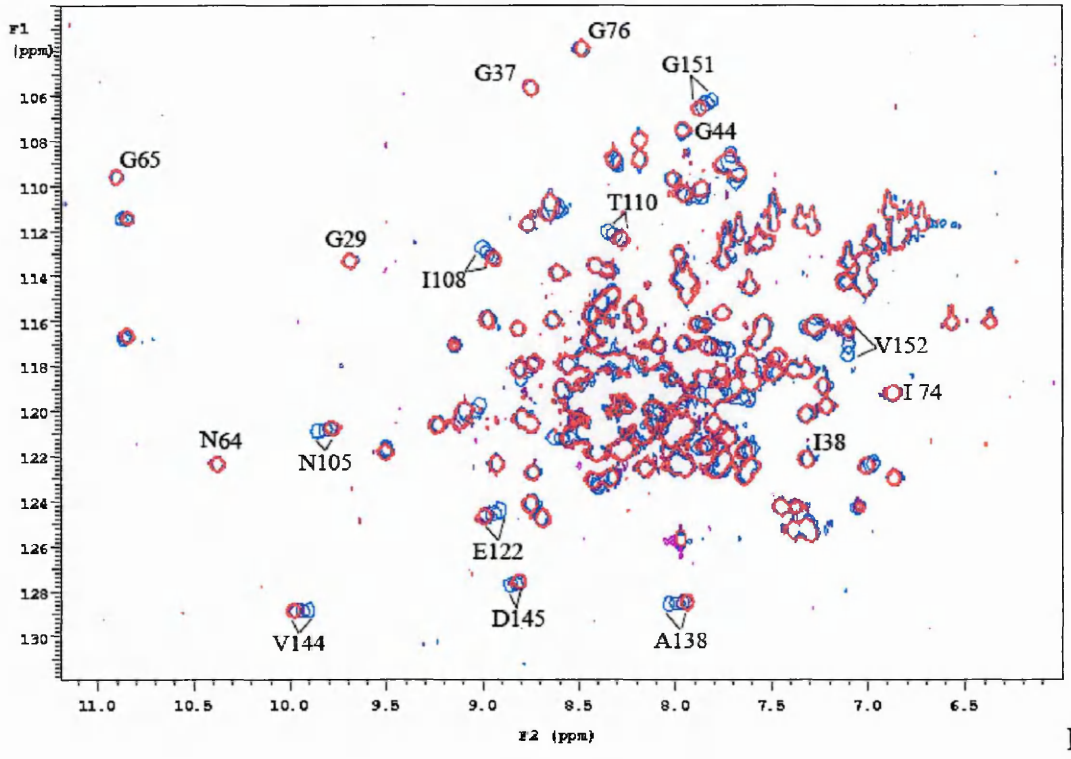


Fig. 5.7.1 Superposition of the HSQC spectra of ^{15}N labelled in the absence (red) and in the presence of a molar excess (~1:3) of peptides (blue). The spectra were recorded at 25°C and 600 MHz. Some of the residues perturbed are labelled.

5.8 Chemical shift mapping for complexes holoF1/TnH(30-61) and holoF1/TnH(126-159)

The observation that only the C-terminal domain signals were perturbed when adding the peptides and that in both cases the peaks perturbed where the same was confirmed by the chemical shift map analysis. Clearly the binding site on F1 C-terminal domain is the same for both peptides.

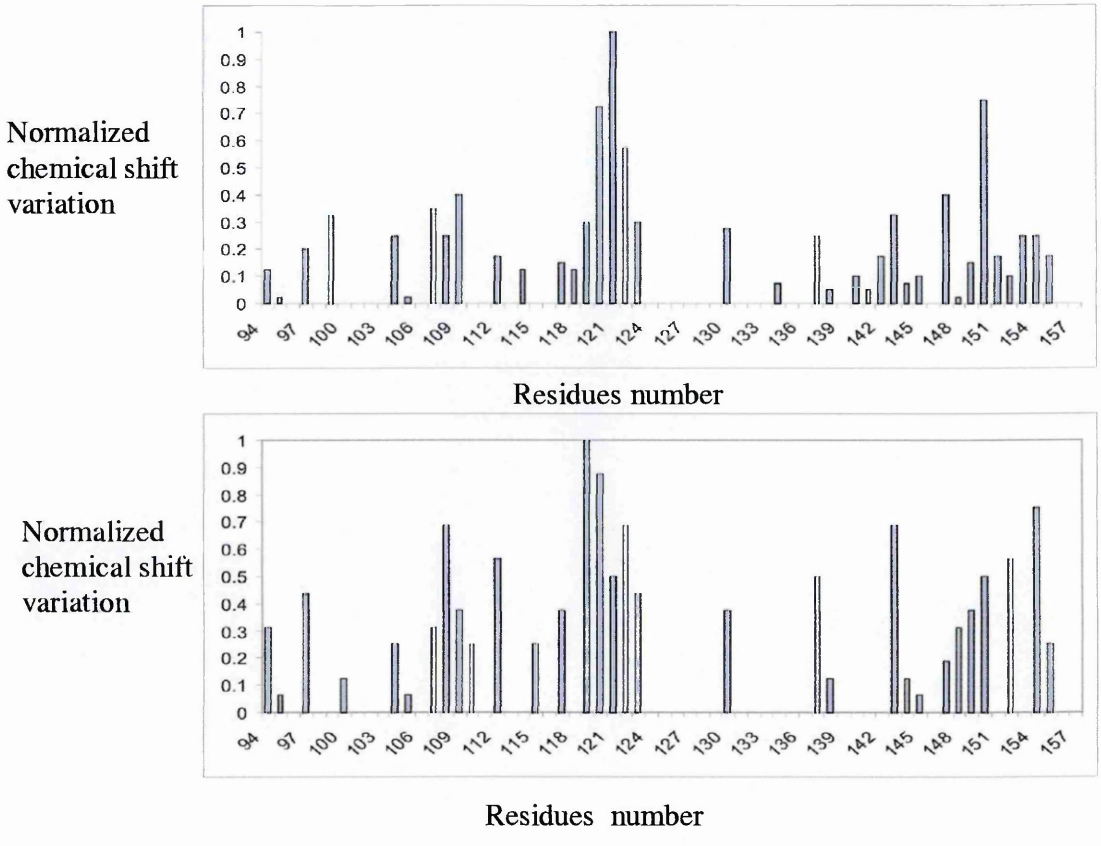


Fig. 5.8.1 Plots of the normalized chemical shift variations induced by addition of TnH(30-61) (bottom) and TnH(126-159) (top) versus the F1 holo TnC sequence. The effects refer to spectra recorded at 1:3 F1:peptide ratios. Only numbering of the C-terminal domain is reported, since no variations are observed for the N-terminal domain. In the F1/TnH(30-61) complex because of the low exchange rate limit, only a tentative assignment could be achieved for residues 101, 114, 116, 117, 127-130, 132-134, 136-137, 147, and 157 which have not shown in the plot. Similarly, residues 97, 114, 127, 129, 130, 132, 133, 136, 147, and 157 could not be precisely identified in the spectrum of the F1/TnH(126-159) complex. The other residues were unambiguously assigned.

5.9 Holo F1, TnH(30-61) and TnH(126-159) native gel

To support the NMR titration data further, we ran a native gel using holo-F1 and each of the two peptides individually or a mixture of the two using different molar ratios of the two peptides. We observed by band-shift complex formation with TnH(30-60). The complex with TnH(126-159) is probably too weak and produces on the gel a smear, as a result the band disappears. The complex formed in co-presence of both peptides runs at the same molecular weight as the one with isolated TnH(30-61), suggesting that TnH(126-159) has no significant affinity for F1 when the C-terminal domain of the protein is blocked with TnH(30-61) and that a second binding site for the TnH(126-159) on the F1 C-terminal domain is not present. These two observations are also confirmed by NMR spectra on the holo-F1/TnH(1-224) and holo-F1/TnH(1-381) complexes.(see chapter 6)

These data therefore do not rule out the possibility that the two complexes could be alternatively formed during muscle contraction and it is possible that competition between two distinct regions of TnH, which alternatively occupy the C-terminal domain of F1, could be at the basis of stretch activation.

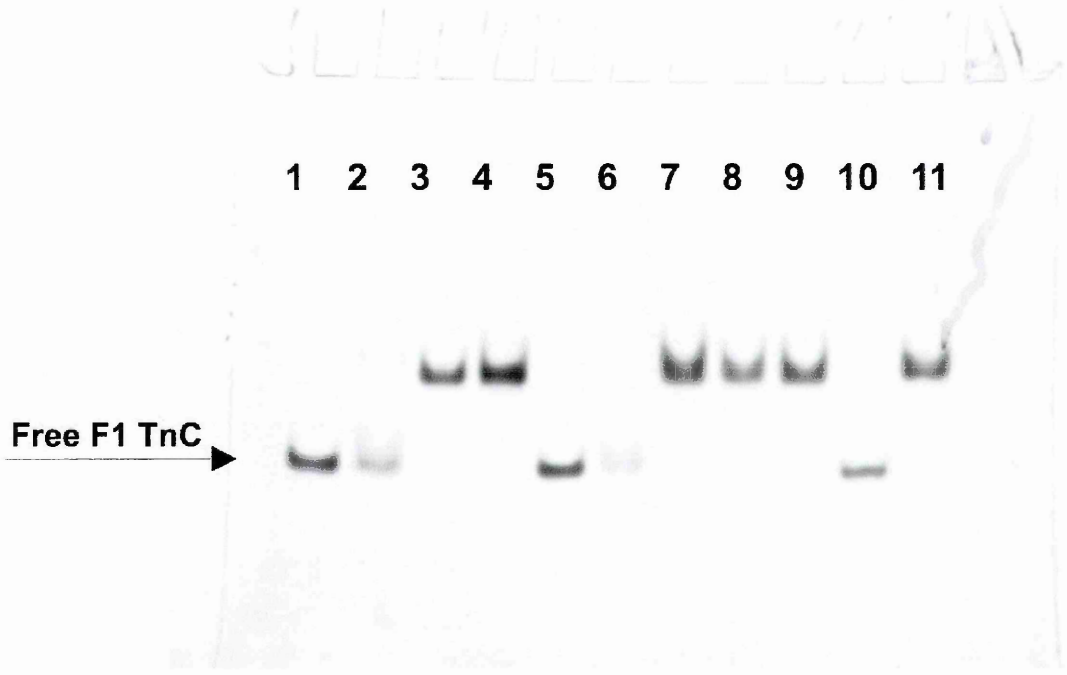
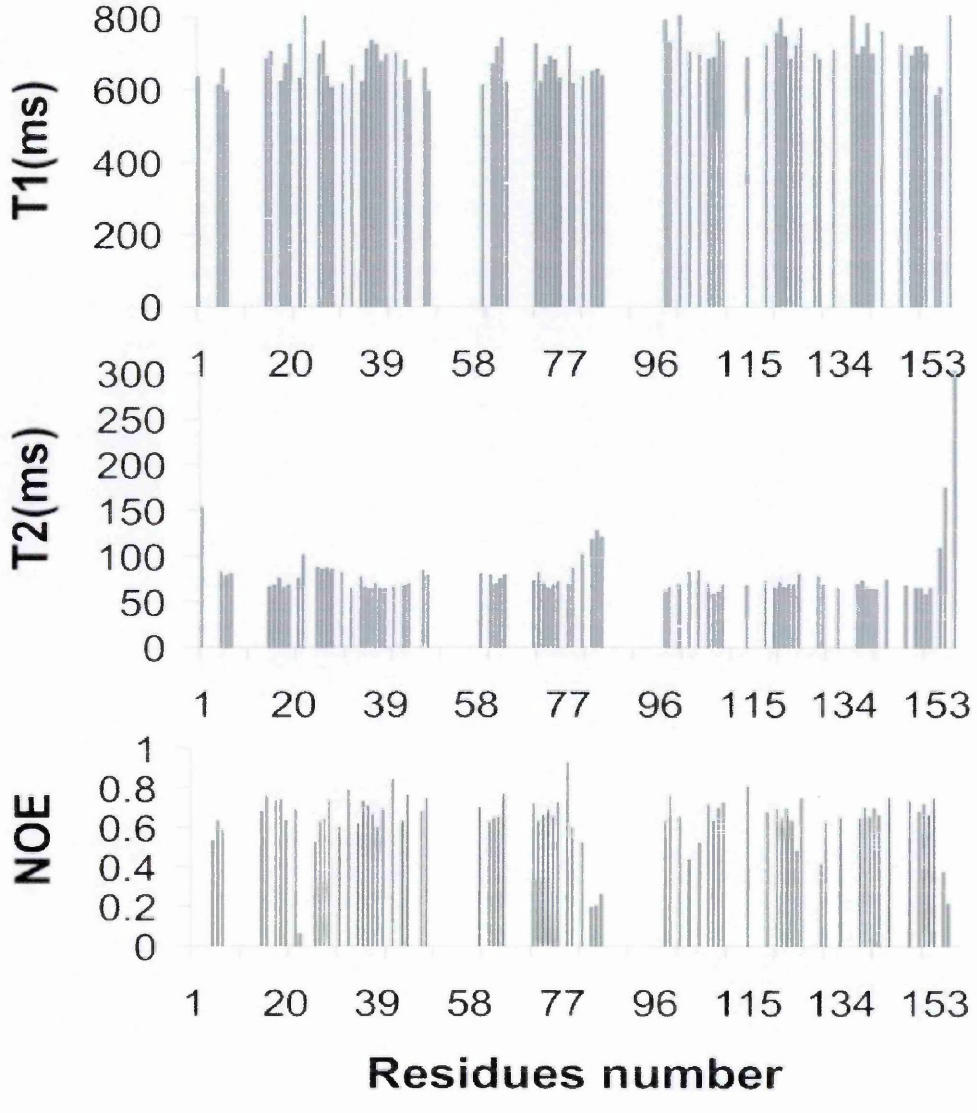


Fig. 5.9.1 Native gel to test the interactions of F1 TnC with the two TnH peptides. The two peptides were incubated individually or together with 22 μM F1 TnC. The lanes correspond as follows: lanes 1, 5 and 10 are controls in which the isolated F1 TnC was used; lane 2: addition of 22 μM of TnH(126-159); lane 3: addition of 22 μM of TnH(30-61); lane 4: addition of an equimolar mixture of the two peptide; lane 6: addition of 66 μM of TnH(126-159); lane 7: addition of 66 μM of TnH(30-61); lane 8: addition of an equimolar mixture of the two peptides. In lane 9, 22 μM of TnH(30-61) and 111 μM of TnH(126-159) were added. Lane 11 is the same as lane 9 but without bromophenol blue.

5.10 Dynamics of holo F1/TnH(30-61) and holo F1/TnH(126-159) complexes

The T_1, T_2 and NOE values were measured for the two complexes For both complexes the T_1 and NOE values increase whereas the T_2 ones decrease upon complex formation. This effect is more evident for the residues belonging to the C-terminal domain. It is also worth noting that there is no substantial difference in the dynamical behaviour between the two complexes. These changes are in accordance with what would be expected for a two domain protein that interacts with a peptide through just one of its domains.



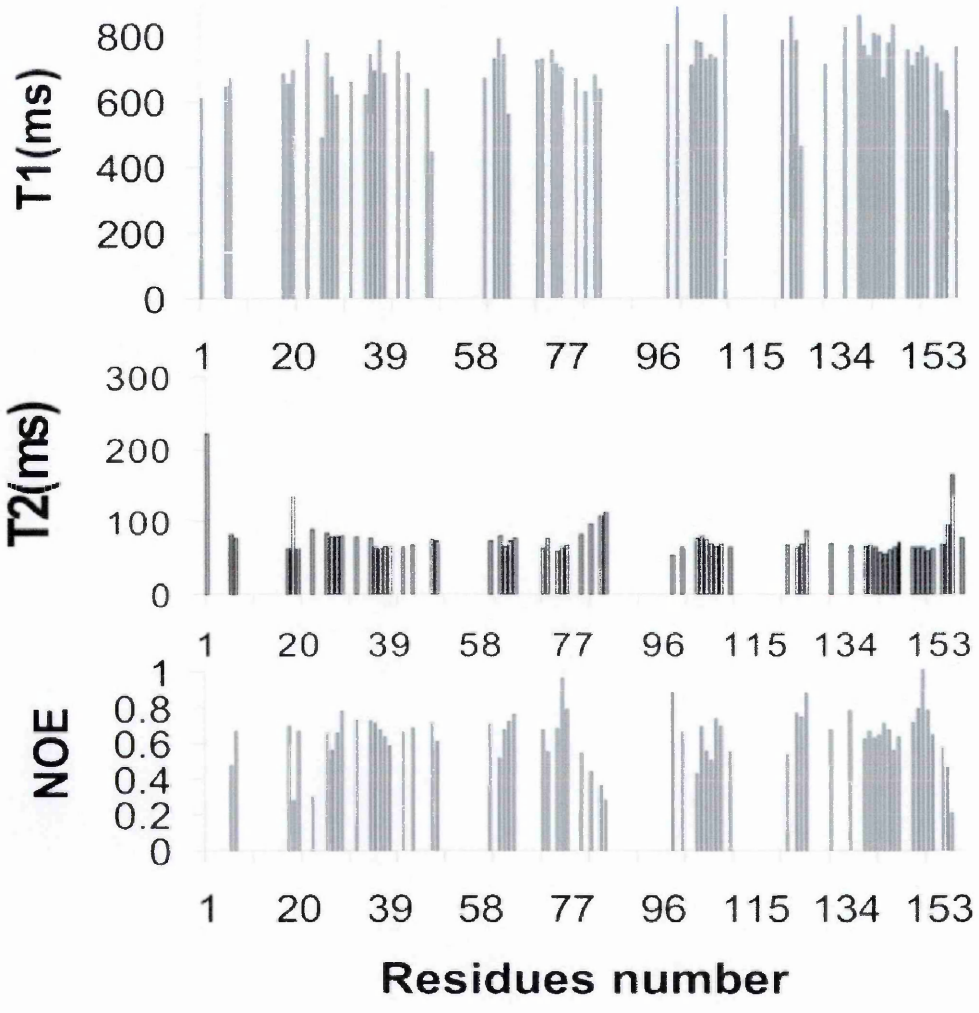


Fig. 5.10.1 Relaxation parameters of the backbone ^{15}N nuclei of holo F1 in complex with TnH(126-159) (top) and TnH(30-61) (bottom) plotted as a function of the residue number. Due to overlapping, poor fitting and lack of assignment, the data for the holo F1 backbone relaxation parameters in complex are incomplete. The data were recorded at 25°C and 500 MHz

5.11 $^{15}N^{13}C$ holo F1 / unlabeled TnH(30-61) complex

As a first approach Haddock (Dominguez, 2003) was used to obtain the structure of the complex. The experimental restraints used during the calculation were the one obtained through the ^{15}N ^{1}H chemical shift perturbation map on the labelled ^{15}N F1.Unfortunately the restraints were not enough for the structures of the complex to converge to a single family. During the many calculation in fact, even though the TnH(30-61) was correctly placed in the proximity of the C-terminal domain of F1, the register and the orientation of the peptide did not converge allowing many possible positions.

In Haddock the residues which are perturbed upon complex formation are defined as "active" or "passive" residues. In the case of NMR titration data, the active residues are defined as residues that show a significant chemical shift perturbation upon complex formation as well as a high solvent accessibility in the free form protein (>50%), the passive ones instead are those residues that show a less significant chemical shift perturbation and/or that are surface neighbours of the active residues and have a high solvent accessibility (>50%). Different lists of active and passive residues have been used during the calculations using less or more stringent thresholds for the solvent accessibility and/or for the chemical shift. In all cases all the peptide residues were considered "active" and its starting structure was assumed to be helical.

Despite many trials none of the calculations as mentioned above converged to a unique cluster. A possible reason for that is the fact that not enough restraints were available for the docking calculation to converge.

In order to get more restraints a different approach was then tried. An $^{15}N^{13}C$ rejected NOESY on a sample containing the complex ^{15}N-^{13}C F1/unlabelled TnH(30-61) at ~0.7mM was acquired. (Zwahlen C., *et al.* 1997) This experiments is designed to filter out all the NOE signals arising from protons bound to ^{15}N and ^{13}C atom. It means that the only visible signals are the NOE arising from protons bound to unlabelled nitrogen or carbon. The signals are resolved in the ^{13}C and ^{15}N dimensions. In our case F1 was labelled and TnH(30-61) was unlabelled, so the NOEs detected were generated by protons belonging to the peptide but, having the F1 labelled, they were resolved in the ^{13}C and ^{15}N plane. In other words the signal x,y and z dimensions were: x the frequency of the proton bound to an unlabeled carbon or nitrogen, (i.e. in this case a proton

belonging to the peptide) z the frequency of either the ^{13}C or ^{15}N (i.e. in this case a carbon or nitrogen belonging to F1) and y the frequency of the diagonal peak (i.e. in this case the proton frequency of the proton bound to the ^{13}C or ^{15}N atom). The diagonal peaks are not visible since is bound to a labelled atom.

Unfortunately most of the NOE peaks from the $^{15}N^{13}C$ rejected NOESY experiment were highly ambiguous because I did not have the assignment of the peptide in the bound form and because the assignment I had for F1 free form was not exactly transferable to F1 in complex. In fact I was able to assign only three unambiguous NOE: CG2 Ile 116, CG1 Val 152 and CG1 Ala 138, the rest of the NOE signals for the above mentioned reason were not assigned. However these three signals are consistent with an homology model of the complex.(see 5.13)

5.12 Unlabeled holo F1/ $^{15}N^{13}C$ TnH(30-61) complex

The unlabeled holo F1/ $^{15}N^{13}C$ TnH(30-61) complex was produced with the aim of solving the high ambiguity of the NOEs found in the previous experiment. An $^{15}N^{13}C$ rejected NOESY was acquired, this time having the TnH(30-61) peptide labelled and F1 unlabelled. The complex was produced cloning the DNA sequence of the peptide in a GST-TEV site containing vector and produced as a fusion tail of GST. Unfortunately due to the sample preparation condition the yield was extremely poor, and the complex produced was barely enough to acquire two $\{^{1}H,^{15}N\}$-HSQC and $\{^{1}H,^{13}C\}$-HSQC spectra. The analysis of these spectra was anyhow extremely informative. In fact analyzing the peptide sequence only two methionines should be present, whereas in the $\{^{1}H,^{13}C\}$-HSQC three C-epsilon methionine signals were visible. These data can be interpreted assuming some kind of conformational exchange affecting the peptide, that happens on a time scale slow enough for the NMR experiment to resolve the two or more species present in solution. The temperature (form 15°C to 35°C) and the ionic strength (from 100mM to 500mM NaCl) were changed in order to minimize the conformational exchange but unfortunately the exchange rate was not affected enough for the appearance of the NMR spectrum to improve (see fig. 5.12.1).

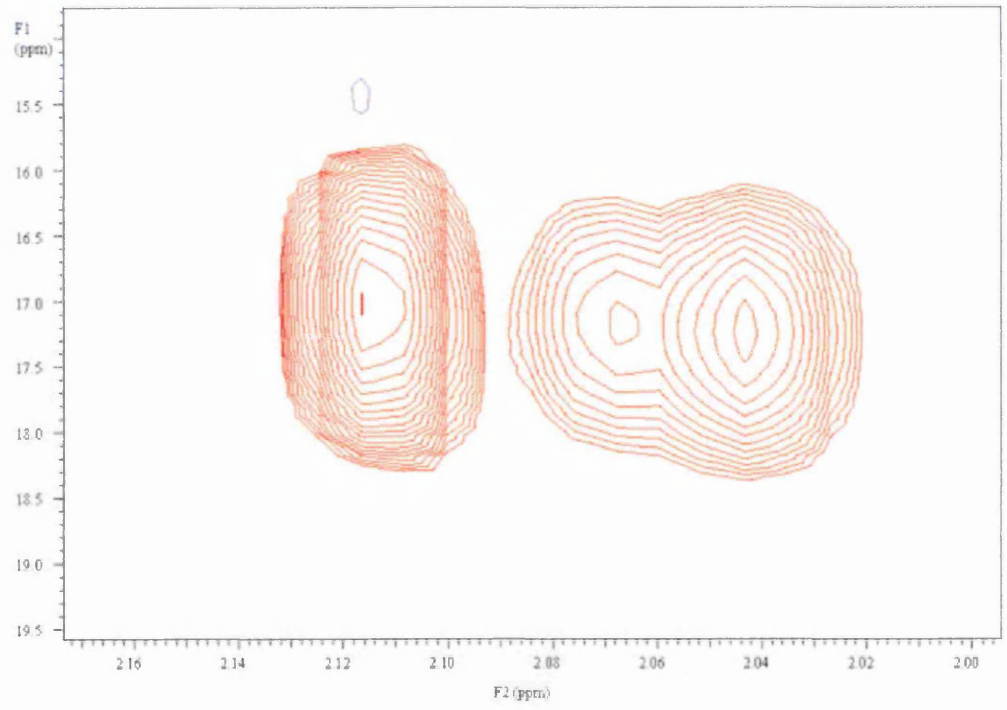


Fig. 5.12.1 The {^{1}H,^{13}C}-HSQC of the ^{15}N and ^{13}C labelled TnH(30-61) peptide in complex with holo F1. The three diagnostic C-epsilon methionine signals are clearly visible

5.13 Holo F1/TnH(30-61) model

Given the experimental difficulties it was decided to build a model of the holo F1/TnH(30-61) complex by homology. The structures of the four TnC/TnI complexes currently available on the C-terminal domain of F1 TnC were superimposed. In addition to the calcium loaded Tn complexes from cardiac and skeletal muscles (Takeda *et al.*, 2003; Vinogradova *et al.*, 2005), the structure of skeletal TnC in complex with the fragment 1-47 of TnI (Vassylyev *et al.*, 1998) and skeletal apo form (Vinogradova *et al.*, 2005) were compared. Their comparison shows that, despite clear differences in the arrangement of the N-terminal domain of TnC and of the interacting region of TnI, the mode of recognition of the C-terminal domain involves the same groove and leads to the same local relative orientation of the TnI and the C-terminal domain of TnC. This

suggests that, independently from the specific sequences involved, the mode of interaction is highly conserved.

When the model was built, using as template the coordinates of the cardiac complex (Takeda *et al.*, 2003), we compared the network of interactions formed with those in known complexes to identified key positions and evaluate their conservation. In cardiac TnI, there are five hydrophobic residues (L53, L54, A56, I57 and L61) which pack against the hydrophobic core of TnC and seem to lead the interaction. Additionally, three positively charged residues of TnI (R45, K46, K50) and E60 form salt bridges with residues of opposite charge in TnC. A similar network of interactions is observed in the skeletal muscle complex (Vinogradova *et al.*, 2005). These contacts are all semi-conserved in *Lethocerus* and lead to a similar network of interactions.(see fig 5.13.1)

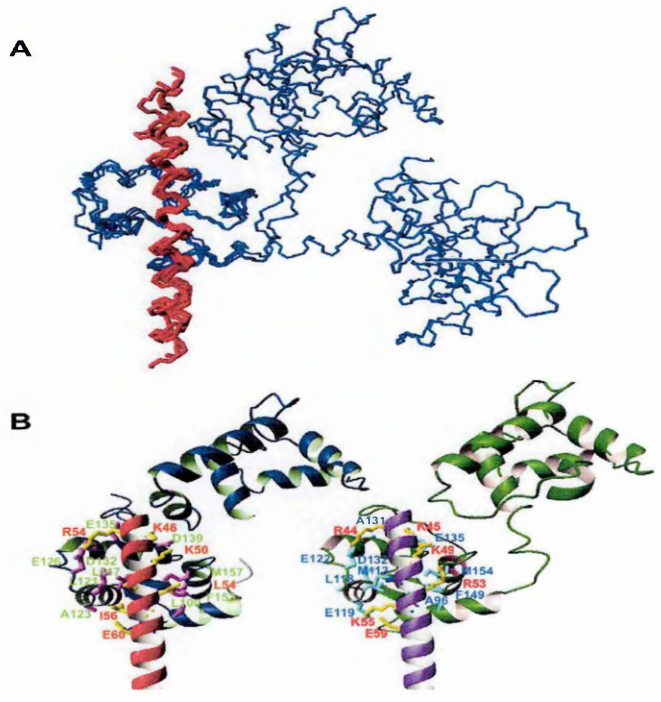


Fig. 5.13.1 Modeling of the complex. A) Structural superposition of the four available structures of TnC complexes (PDB entry:1a2x, 1j1d, 1ytz and 1yv0) with the N terminus of TnI. The structures were superposed on the C-terminal domain of TnC. The comparison shows that, despite the differences in the relative orientation of the N-terminal domain, TnI interacts in the same groove and adopts the same relative orientation to the C-terminal domain. B) Comparison of the network of interactions formed between the C-terminal domain of TnC and TnI in the cardiac complex (Takeda *et al.,* 2003) (left panel) and in a homology model built using the cardiac structure (1j1d) as a template. The side chains of the key residues involved are displayed explicitly and labeled in red and green according to whether they correspond to TnI and TnC respectively, for the cardiac structure. Similarly, they are coloured in red and blue according to whether they correspond to *Lethocerus* TnH and TnC respectively (right panel).

Chapter 6

RESULTS: Result: F1, TnH and TnT interaction

6. Result: F1, TnH and TnT interaction

6.1 Holo F1/TnH(1-224) and holo F1/TnH(1-381) complexes

Given the low sequence similarity of the *Lethocerus* and vertebrate Tn sequences, it is in principle possible that other regions of TnH outside those tested could establish interactions with the N-terminal domain of F1 To clarify this point two longer clones of TnH: TnH(1-224) and TnH (1-381) spanning respectively residues 1-224 and 1-381 were produced. These constructs include both peptides previously tested and exclude only the low complexity C-terminal tail. Because of its highly repetitive character, it is unlikely that this region could contribute to specific binding with F1 TnC. The complex was formed dissolving holo F1 and the TnH construct in 6 M urea and then dialyzing the urea out (see Chapter 8). The $\{^{1}H,^{15}N\}$-HSQC-NMR spectrum of the complexes F1/TnI(1-224) and F1/TnI(1-381) were then acquired. Taking into account that the broadening of the peaks is due to the fact that the complexes have a bigger molecular weight and tumble on a lower time scale, the complexes HSQC spectra can be almost overlapped to the HSQC-spectra of the complex F1-TnH(30-61). This means that the only portion of the longer clones interacting is TnH(30-61), that there is no second binding site on the F1 C-terminal domain for TnH(126-159) and that the N-terminal domain of F1 does not interact with TnH.

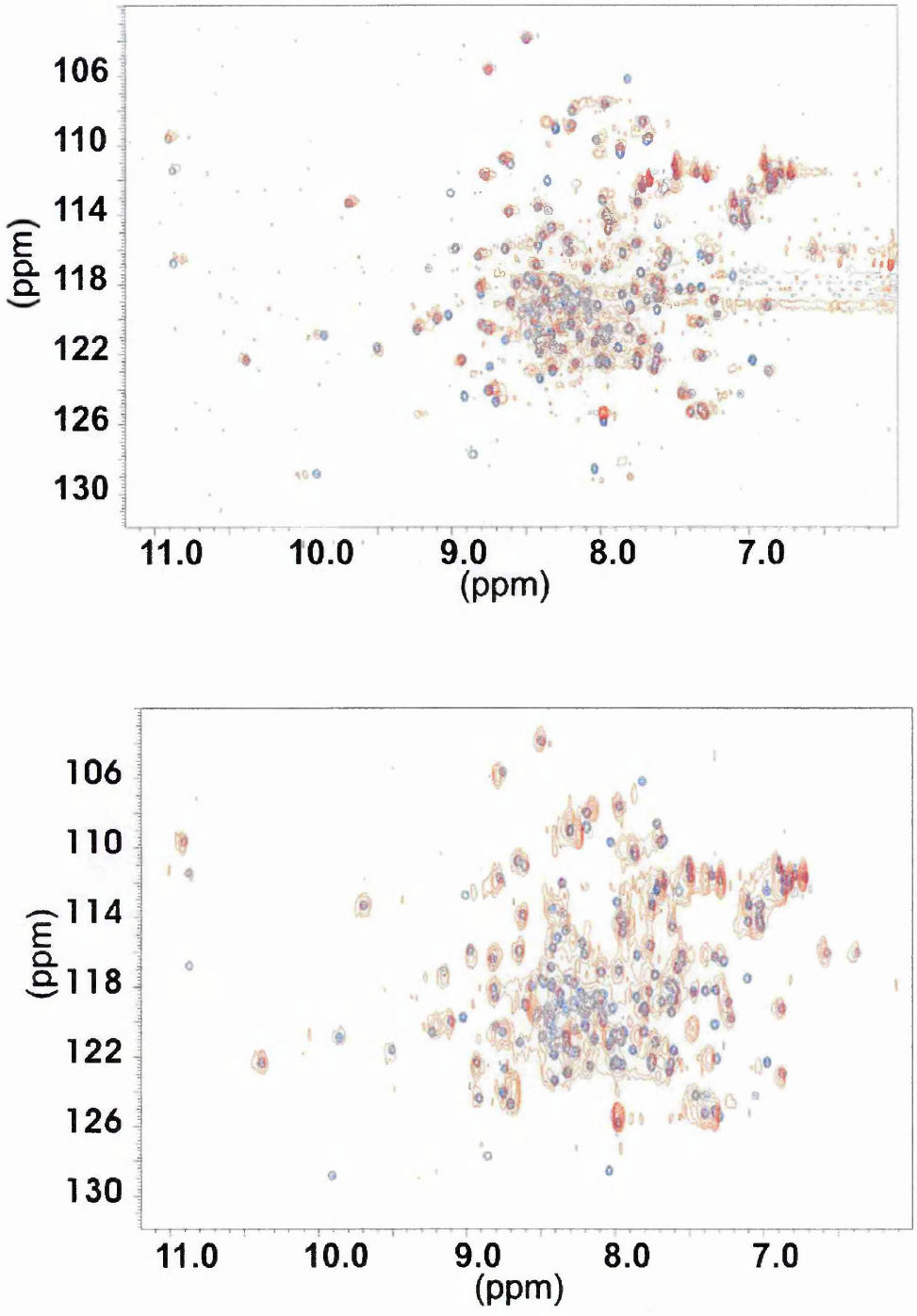


Fig. 6.1.1 Searching for regions which could interact with the N-terminal domain of holo F1. HSQC spectra of the holo F1/TnH(1-224) (top) and holo F1/TnH(1-340) (bottom) complexes. In both spectra the free F1 peaks are in blue whereas the complex ones is red.

6.2 Holo F1/TnT(110-266) complex

To check if the N-terminal domain of F1 interacts with TnT a clone TnT(110-266) spanning residues 110-266 of *Lethocerus* TnT was produced by B. Bullard and tested by me by NMR. The complex was made in the same way as for the TnH constructs and an $\{^{1}H,^{15}N\}$-HSQC spectrum was acquired. Minor chemical shift perturbation was observed suggesting that at least in these conditions F1 and TnT do not interact.

6.3 *Lethocerus* holo Tn Trimer

To rule out the possibility that TnT was needed for the interaction between TnH and the N-terminal domain of F1 to happen we decided to reconstruct the whole *Lethocerus* troponin trimer. The trimer was produced first dissolving all the component in 6M urea and then dialyzing it out gradually. The $\{^{1}H, ^{15}N\}$-HSQC spectrum was then acquired. After analyzing the spectra, only the cross peaks belonging to the N-terminal domain of F1 are visible and they all resonate at the same chemical shift frequencies as the free holo F1, whereas all the cross peaks belonging to the C-terminal domain have disappeared. Given the size of the trimer (52KDa) and the fact that no deuteration was used to decrease the rate of transverse magnetization decay is remarkable that almost all the F1 N-terminal domain signal are visible. These data suggest that in the *Lethocerus* troponin trimer the N-terminal domain of F1 is able to tumble almost independently from the rest of the complex, behaving at least from an NMR point of view as an independent domain. It is therefore safe to assume that it does not interact with any of the other subunits. In principle is not possible to rule out the possibility that the portion of TnH and TnT not tested could interact with the N-terminal domain of F1.

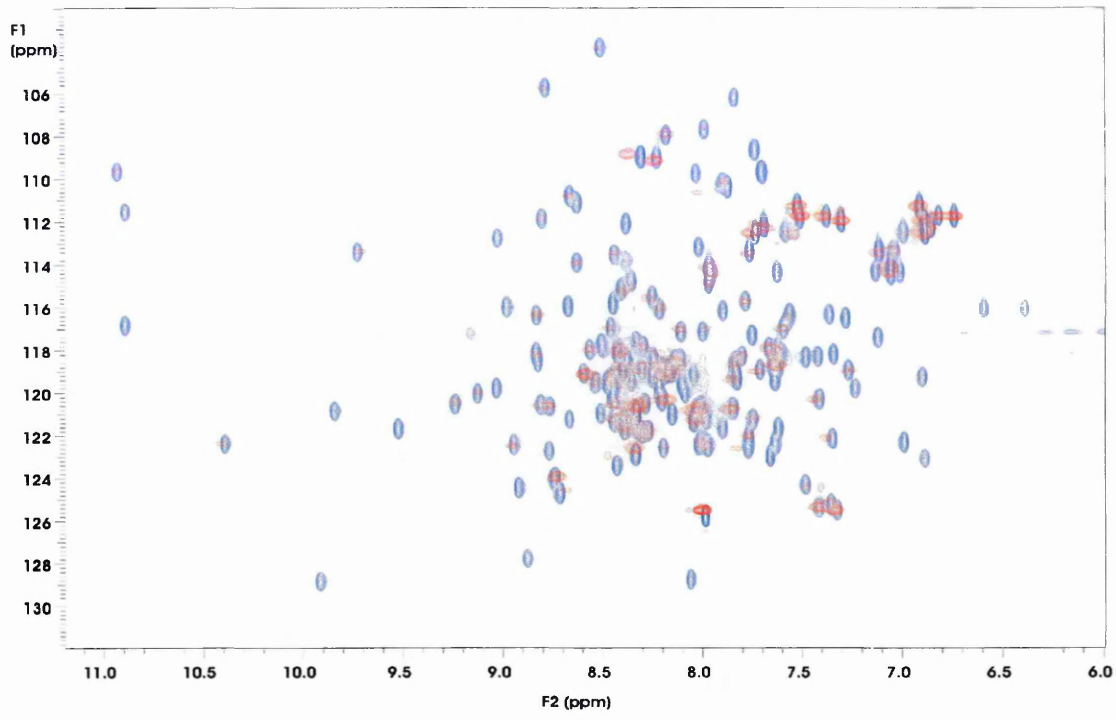


Fig. 6.3.1 Superimposition of the {^{1}H-^{15}N} HSQC-spectra of holoF1(blue) and the holo F1/TnH(1-224)/TnT(110-266) complex (red). In the complex the only visible signals belong to N-terminal domain of holo F1 and they almost exactly overlap to the correspondent signal in the free form. This observation confirms that in such condition the N-terminal domain of F1 does not interact with any other portion of the troponin trimer

CHAPTER 7

DISCUSSION

7 Discussion

7.1 Summary of the result presented in this thesis

The structure of F1 provides the first description of one of the major components that is expected to determine the regulation of asynchronous muscle contraction. In spite of the obvious similarities with other TnC structures, the *Lethocerus* protein has some specific features, which may be related to its function. F1 does not contain the N helix, having a shorter N-terminus compared to other TnC sequences. This helix has been shown to increase Ca^{2+} affinity (Chandra et al., 1994). It is therefore unnecessary in the Ca^{2+} insensitive N-terminal domain of F1. The Ca^{2+} insensitive N-terminal domain is in a closed conformation. This is in agreement with what is observed in other TnCs in the absence of Ca^{2+} and, more generally, in apo-EF-hand domains (Gagné et al., 1998; Yap et al., 1999). More unusually, the C-terminal domain of F1 TnC is in an open conformation in the presence of only one Ca^{2+} ion. This is at variance with the N-terminal domain of cardiac TnC, which also binds only one Ca^{2+} ion but is in a closed conformation regardless of whether it is in its calcium-free or calcium-loaded forms. In cardiac muscles, Ca^{2+} binding does not induce a significant change in the structure of the protein. (Sia et al., 1997) The presence of both Ca^{2+} and TnI are necessary to stabilize an open conformation, which leads to exposure of the hydrophobic pocket. (Li et al., 1999)

The dynamic behaviour and the stability of Ca^{2+}-loaded F1 were studied. No substantial difference in protein flexibility within each of the two domains is observed, suggesting that they have similar dynamic properties. As in other TnC structures, the linker is highly flexible. This region clearly constitutes a hinge which makes the two domains structurally semi-independent. The N and the holo C-terminal domains have comparable thermal stability.

Apo-F1 structure and thermal stability were also studied. In apo F1 the N-terminal domain is significantly more stable than the apo C-terminal domain. The apo C-terminal domain despite being structured shows an appreciably lower stability than the holo C-terminal domain. This feature is also observed in skeletal TnC and in calmodulin. (Fredricksen and Swenson, 1996; Masino et al., 2000) The apo C-terminal domain of F1 TnC is, however, more stable than the corresponding region in skeletal TnC, which is

unfolded at room temperature. (Mercier et al., 2000) It is also important to note the substantial structural identity between the holo and apo F1. The F1 apo C-terminal domain presents an open conformation also in the absence of calcium. This feature is also seen with calbindin D9k, a calcium binding EF hand protein that is an open conformation in both the apo and holo forms.(Kordel *et al.*, 1993, Skelton *et al.*, 1995)

Finally the interactions of F1 TnC with TnH and TnT using different constructs and complementary techniques have been studied. Only the C-terminal domain binds TnH in the binary complex, and the F1 N-terminal domain appears not to interact with any of the subunits of the troponin trimer. This finding indicates that stretch activation is regulated by a mechanism that is completely different from that observed in Ca^{2+}-activated muscles, and suggests that regulation is achieved through the C-terminal domain rather than the N-terminal domain. A similar hypothesis was proposed for scallop troponins. (Ojima et al., 2000)

The C-terminal domain binds both the N-terminus of TnH and a peptide spanning the inhibitory and switch regions.

The complexes with the two peptides differ greatly in their affinities. The interaction with the TnH N-terminus is in the nanomolar range, which is comparable with the value reported for the skeletal TnC/TnI complex. (Tripet et al., 2003) The K_d of the TnC/TnH(126-159) complex is in the micromolar range, as in the complex of skeletal TnI(96-111) with the C-terminal domain of TnC. (Mercier et al., 2000) Could the two complexes be alternately formed during muscle contraction? A model in which two regions of TnI compete for the C-terminal domain of TnC has already been proposed for skeletal muscle. (Tripet et al., 1997) This hypothesis was however subsequently discredited, by the identification of a second site of weaker interaction on the C-terminal domain and by structure determination of the Tn complex. (Vinogradova et al., 2005) To check whether a similar situation could occur in Lethocerus, and to further confirm the NMR data on the complexes F1/TnH(1-224) and F1/TnH(1-381), the experiment proposed by Tripet et al. (2003) based on a native gel was repeated. An additional band-shift in the presence of both peptides, as expected if both interactions were possible at the same time is not observed. It therefore remains possible that competition between two

distinct regions of TnH, which occupy the C-terminal domain of TnC alternately, could be the basis of stretch activation.

7.2 A simple model to explain stretch activated contraction

Here I suggest a mechanism to explain the data. Although still not completely clear , this mechanism could be, in its general components, the basis for stretch activation. Some other component not analyzed in this thesis could also play a role in stretch activation.

A simple model could be made to explain stretch activation based on the results of this thesis. This model is based on the observation that in the experimental conditions tried the N-terminal domain of F1 does not interact with any of the troponin subunits, and on the observation that the C-terminal domain of F1 can bind with different affinity to two different portions of TnH: to TnH(30-61) in a calcium insensitive way and to TnH(126-159) in a slightly calcium sensitive way.

The last statement: the C-terminal domain of F1 can bind to TnH(126-159) in a calcium sensitive way is based on the experimental evidence that upon calcium binding to the C-terminal domain of F1 the binding affinity towards TnH(126-159) is increased by a factor of 1.6 at 20°C (a factor of 2.4 at 40°C). This difference in affinity between the holo and apo forms is within the experimental error of the biophysical techniques used to measure the affinities. The model presented here is based on this small difference and I will try to make a convincing case for the fact that such a difference has a physiological meaning especially when considering the whole muscle fibre.

The data as a whole suggests that in stretch activation the regulation is achieved in a completely different way compared to vertebrates. The general accepted view in vertebrates, that implies a structural role of the C-terminal domain of TnC and a regulatory role for the N-terminal domain of TnC, can not be transferred to IFM fibres.

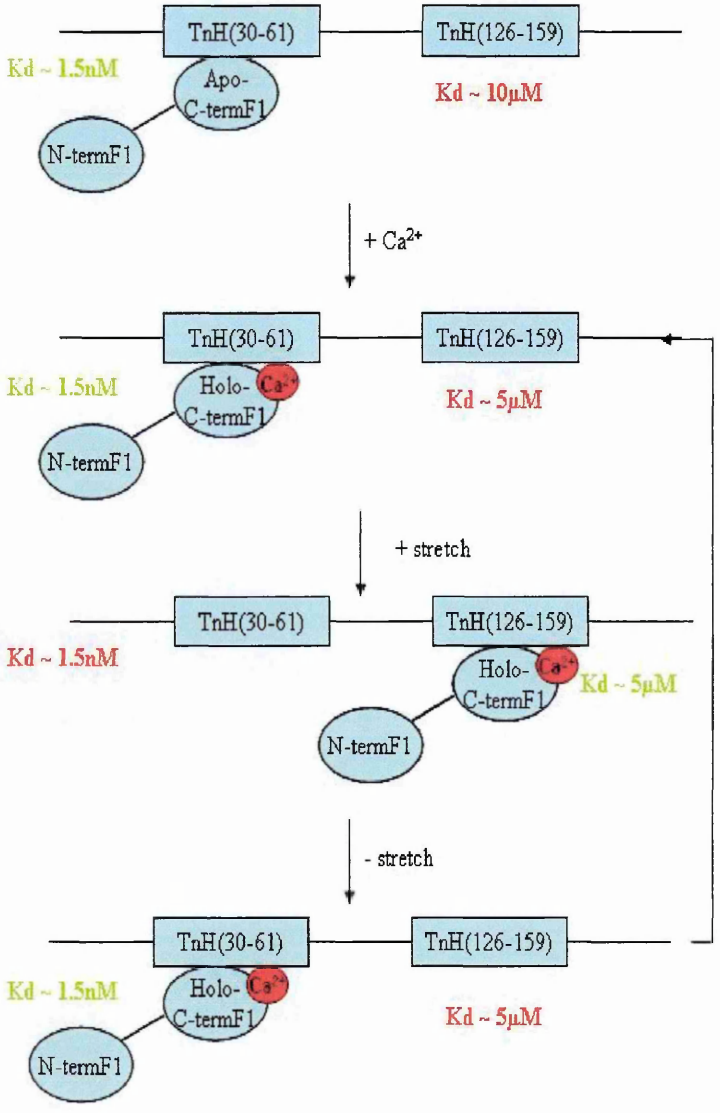
Bullard and co-workers have demonstrated that the presence of calcium is crucial for stretch activation in IFM fibres and the general belief was that, in stretch activated fibres, calcium was needed in order to induce an open conformation in the F1 C-terminal domain which would cause F1 to bind TnH. According to this hypothesis, calcium and the C-terminal domain of F1 would have only a structural role leaving open the question of how regulation could be achieved. The experiments carried out in this thesis rule out

this hypothesis: calcium is not needed for the interaction between F1 and TnH and therefore for the anchoring to the thin filament. The affinities between TnH(30-61) and F1 are exactly the same in the presence and in the absence of calcium. Moreover the C-terminal domain of F1 in the apo form is structured and already in an open conformation. Calcium binding does not induce any major structural change.

A possible alternative explanation for the role of calcium in stretch activation is that calcium has instead a regulatory role but not on the N-terminal domain like in vertebrates but on the C-terminal domain. The regulation is achieved by modulating the TnH(126-159) peptide binding affinities. In this model the C-terminal domain of F1 would have at the same time a regulatory (i.e. binding to TnH(126-159) in a calcium dependent manner) and structural role (i.e. binding to TnH(30-61) in a calcium independent manner).

Such a model would not need the pumping out of calcium for muscle relaxation making deactivation extremely fast. It would also explain the role of calcium, which is that of lowering the difference in affinity between the two TnH binding sites. It could be argued that the binding affinity difference to TnH(126-159) between apo F1 and holo F1 is extremely low but it has to be considered that in *Lethocerus* IFM fibres every single myosin filament is surrounded by three actin filaments and that every tropomyosin molecule is associated with a troponin subunit. Considering that *Lethocerus* IFM sarcomere is 2.5μm long, that the actin filament with tropomyosin and troponin is 1.15μm long, that there are two tropomyosin-troponin complexes per 38.0nm along the actin filament, and finally that the ratio F1:F2 is 7:1 is possible to calculate the number of F1-TnH couples per each three actins associated with each myosin filament. That number is 160, and the small calcium dependence would apply for every couple F1/TnH(126-159). In other words the difference in affinity between HoloC-termF1-TnH(126-159) and ApoC-termF1-TnH(126-159) might be not significant on a single couple F1/TnH(126-159) but it might became significant when the whole actin/myosin bundle is considered This model would also explain why the inhibitory peptide is conserved across vertebrates and insect whereas the switch peptide is not. This feature suggests that the inhibition of muscle contraction is achieved in a similar way in vertebrates and insects: through the binding of the inhibitory peptide on the actin surface blocking the actin-myosin

interaction whereas the regulation of muscle contraction is achieved in different ways: through the binding of the switch peptide to the N-terminal domain of TnC in a calcium dependent manner in vertebrates and through the sliding of TnH into the C-terminal domain of F1 in insects.



Step 1: the system is calcium free. F1 is in its apo form the difference in binding affinity causes TnH(30-61) to bind to apo C-term F1

Step 2: calcium is released from the sarcoplasmatic reticulum into the myofibril. Now the Holo-F1 bindin affinity for the TnH(126-159) is higher than the Apo form but still the difference in affinity is such that TnH(30-61) stays bound to holo C-term F1

Step 3: somehow stretch causes TnH(30-61) to detach from Holo C-term F1 and puts the C-term F1 lobe in the vicinity of TnH(126-159). Inhibition is released and the cross-bridges formation can start

Step 4: stretch is released muscle can relax and the system is back to step2. a second contraction cycle car start

Scheme 7.2.1 A schematic representation of how stretch activation could work in IFM That model would not require the pumping out of calcium for deactivation to happen. As soon as stretch is released a new cycle can start.

It is also worth noticing that in *Lethocerus* stretch activation is temperature dependent. In the wild, it happens at temperatures above 40°C: the *Lethocerus* needs to warm up before it can fly. It is temping to speculate that this temperature correlates with the unfolding temperature of the apo C-terminal domain of full length F1 (~36°C). This temperature dependence flight behaviour is the reason why some of the peptide binding assay were repeated at 40°C. The difference in affinity of F1 apo and holo for the TnH(126-159) increases with the temperature.

7.3 Further work to validate this model

Here I am going to suggest some experiments that could be done to validate this model. First the peptide TnH(126-159) comprises both the equivalent of the inhibitory and the switch peptide in vertebrates. It could be therefore worth to check the interaction with F1 using a peptide which spans only one of the two segments. In this way it could be checked if the regulation is achieved through direct competition for the inhibitory peptide between the C-terminal domain of F1 and actin or as in vertebrates through a general structural tropomyosin-actin rearrangement caused by the binding of the switch peptide to the C-terminal domain of F1.

A second experiment I would suggest is to repeat the stretch activation on IFM skinned fibres at 40°C to reproduce the in vivo condition. At that temperature the calcium dependent affinity difference of the C terminal domain of F1 towards the TnH(126-159) peptide is higher and therefore the force applied to stretch activate such a system should be lower.

Finally according to this model the N-terminal domain would be left without any role, as a left over of evolution. A way to prove the latter statement would be to test the activity of skinned fibres where the F1 subunit in the troponin trimer is substituted with a subunit containing just the C-terminal domain without the N one. If such fibres would show the same activity as the ones containing F1 full length it would be a good evidence to prove that the N terminal domain of F1 is a left over from evolution without any structural nor regulatory role.

7.4 Could F1 just be the mirror image of half of cardiac TnC?

Finally I would like to draw some attention into the many similarities present between cardiac and IFM troponins: both TnC subunit have a single calcium binding domain that does not change conformation upon calcium binding, in both cases calcium causes a difference in affinity towards the switch peptide without causing major structural changes, both TnI subunits have an extension (N terminal in cardiac C terminal in IFM) that can be phosphorylated and both muscle fibres exhibit a stretch activated contraction mechanism. (Li *et al.*, 2003, Ward *et al.*, 2004, Mateos *et al.*, 2005)

If the similarities and differences between cardiac TnC and F1 TnC are analyzed, it is tempting to speculate about the way calcium and stretch regulation is achieved in cardiac muscle. It could be possible to consider cardiac TnC as the sum of F1 and F2. In other words whereas in IFM fibres the F1 isoform regulates stretch activated contraction and the F2 isoform regulates the calcium activated one, in cardiac fibres just one TnC isoform is enough to regulate both types of contraction: the calcium regulated one using both lobes and the stretch activated one using just the N-terminal domain: the domain that has some similarities with the C-terminal domain of F1. It is as if cardiac TnC turns into F1 in a mirrored fashion performing the equivalent of F1 function through its N-terminal domain and inactivating its C-terminal domain. The scenario described above it is just a hypothesis but it has to be noted that the presence of stretch activated regulation in heart has been demonstrated. (Steiger 1971, Steiger 1977, Vemuri *et al.*, 1999) What also is still not clear is the exact role of such mechanism in cardiac fibres. Do the two mechanisms (calcium regulated and stretch activated) coexist at the same time or are cardiac fibres stretch regulated only when the heart rate rises above a certain threshold and if this is the case how is the switch achieved, could the TnI phosphorylation have a role in it. These are just some of the many questions that need to be answered to have a deeper understanding of stretch regulated muscle contraction in hearth, IFM and muscles in general.

Chapter 8

MATERIALS

8 Materials

8.1 F1 plasmid production

The *Lethocerus* F1 cDNA inserted into the *NcoI/NotI* sites of a modified pET24d (M11) expression vector (Novagen) containing an N-terminal hexahistidine (His_6) tag followed by a tobacco-etch-virus (TEV)-protease cleavage site was kindly provided by Dr B. Bullard.

8.1.1 Transformation of competent cells

The construct were transformed through a heat shock protocol using BL21(DE3) or BL21(DE3) pLysS competent cells (Stratagene). The plasmid DNA (1μl at 50-100 ng/ μl) was added to 50 μl of competent cells. The cells were kept on ice for 30 minutes and then heat shocked at 42 °C for 30 seconds followed by incubation on ice for 2 minutes; 300 μl of Luria-Bertani broth (LB) were added to the cells, left at 37 for ~1h and then spread on LB agar plates containing kanamycin (30 μg/ml).

8.1.2 Fermentation

Recombinant unlabelled protein was grown in LB broth containing 30 μg/ml of kanamycin. A single colony from the LB agar plate was used to inoculate 100 ml per litre of media. The inoculation was left to grow at 37 °C until the optical density measured at 600nm (OD_{600}) was around one. The 100 ml growth was then used to inoculate one litre of media and the cell grown at 37 °C. When OD_{600} was ~0.7-0.8 isopropyl β-D-thiogalactopyranoside (IPTG) at 0.5mM final concentration was added to the cell to induce the protein production. After induction the cells were left to grow for ~3-4 h.
Isotopic labelled F1 was produced growing the cells in minimum medium supplemented with [$^{13}C_6$] glucose and/or $^{15}NH_4Cl$ as sole source of carbon and nitrogen respectively.

8.1.3 Cells lysis

F1 is expressed in the cytoplasm. Cells were centrifuged at10,000 rpm for 5 minutes at 4° C and the pellet re-suspended in 10 ml lysis buffer per litre of media. The lysis buffer contains 150mM NaCL, 5mM imidazole, 20mM Tris, lysozyme, DNAse I and a tablet of 'Complete Inhibitor Cocktail'. (Roche) The lysate was sonicated for 20 minutes on ice

using a 50% duty cycle with a 50 watt output. The supernatant was collected after centrifugation at 20,000 rpm for 40 minutes at 4° C.

8.1.4 Nickel NTA agarose protein purification

The His_6 tagged F1 was purified from the cell extract using Ni^{2+}-nitrilotriacetate (Ni-NTA)–agarose columns (Qiagen). A column was packed with Ni-NTA (1ml per litre of growth of medium) and equilibrated with five column volumes of lysis buffer. The cell lysate was run through the column twice, then the column was washed twice each time with a low salt and high salt buffer. The composition of the low salt salt is 50 mM NaCl 20mM imidazole, the high salt one is 1M NaCl, 20mM imidazole. These washes are designed to remove any non specific bound contaminants from the column. F1 was eluted using an elution buffer containing 100mM NaCl and 300mM imidazole. All the fraction were checked on a 10% acrylamide gel.

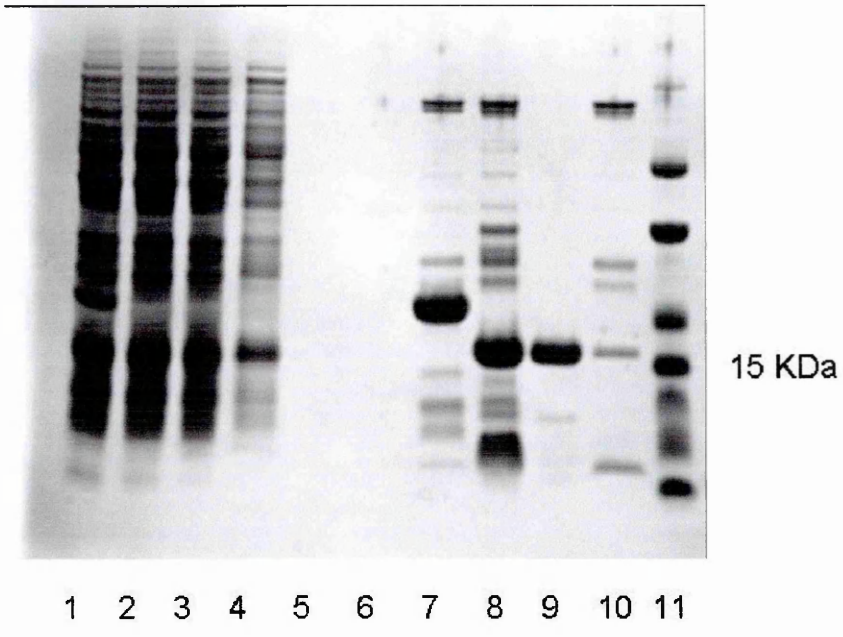


Fig. 8.1.4.1 10% acrylamide gel of the F1 TnC purification fractions. The lanes correspond as follows: lane1: lysed cell supernatant, lane 2: 1st load flow-through, lane 3: 2nd load flow-through, lane 4:low salt buffer 1st wash, lane 5: high salt buffer wash, lane 6: low salt buffer 2nd wash, lane 7: 300mM imidazole eluate, lane 8: Imidazole eluate after incubation with TEV, lane 9: TnC eluate, lane 10: 2nd 300mM imidazole wash, lane 11:molecular markers.

8.1.5 TEV cleavage

The His_6 tag was removed using TEV protease. TEV is a specific protease which recognise the seven residues long tag(GLU-X-X-TRY-X-GLN-SER). (Phan *et al.*, 2002) The cleavage happens between GLN-SER. 2-3 μl of TEV at 25 μM were used per litre of media. The reaction was left at room temperature for 16 hours.

8.1.6 Gel Filtration protein purification

Size exclusion chromatography was used when necessary to remove the cleaved His_6 tag and to further purify F1 from contaminants. The reaction mixture was concentrated to approximately 5ml using a filter concentration (Amicon) and then loaded on a Superdex 75 16/60 column (Pharmacia). The column was equilibrated with a buffer containing 100mM NaCl, 20 mM Tris 0.2% NaN_3 at pH 6.8. The buffer was de-gassed and filtered using 0.25μM filter. The elution profile was then analyzed and the fraction of interest screened using 10% acrylamide gel.

8.1.7 Buffer Exchange using PD-10 columns

F1 samples when need were buffer exchanged using Sephadex G-25M PD-10 column (Pharmacia). The protocol suggested by the manufacturer was followed: first the column was equilibrated with the chosen buffer, then 2.5ml of ~1mM F1 solution was loaded on the column and finally the F1 was eluted with 3.5 ml of the chosen buffer.

8.1.8 Concentration Determination

The Beer-Lambert law (eq.8.1.8.1) was used to measure the protein concentration. A double beam HELIOS α spectrophotometer (UNICAM) was used for the measurement and the extinction coefficient was calculated using eq.2 (Pace *et al.*, 1995)

$$Abs = \varepsilon * c * l$$

Eq. 8.1.8.1 Beer-Lambert law, ε is the extinction coefficient, l is the path length in cm and c is the molar concentration

$$\varepsilon_{280} = (\text{Trp})*(5685) + (\text{Tyr})*(1285) + (\text{Cys})*(125)$$

Eq. 8.1.8.2 Calculated extinction coefficient for an unfolded protein in guanidine hydrochloride

8.2 C-terminus F1 expression and purification

To clone the C-terminus of F1 two set of primers were designed based on the cDNA sequence of F1 full length. The sense primer sequence is: 5´-CATGGCAGCAGGAACTGCGCGAAGCGTTTCGCCTGTATT-3´ the anti sense one is 5´-GGCGGAATTTATGGGCGTGATGGGCGTGATGACCGGCGG-3´.

The PCR product was gel purified using the Novagen bla bla kit, following the protocol suggested by the manufacturer. The PCR purified product was then digested using NcoI and NotI restriction enzymes. The product was then cloned into the NcoI/NotI sites of a modified pET24d (M11) expression vector (Novagen) containing an N-terminal hexahistidine (His_6) tag followed by a tobacco-etch-virus (TEV)-protease cleavage site. The expression and purification procedures are the same as the one used for the full length F1.

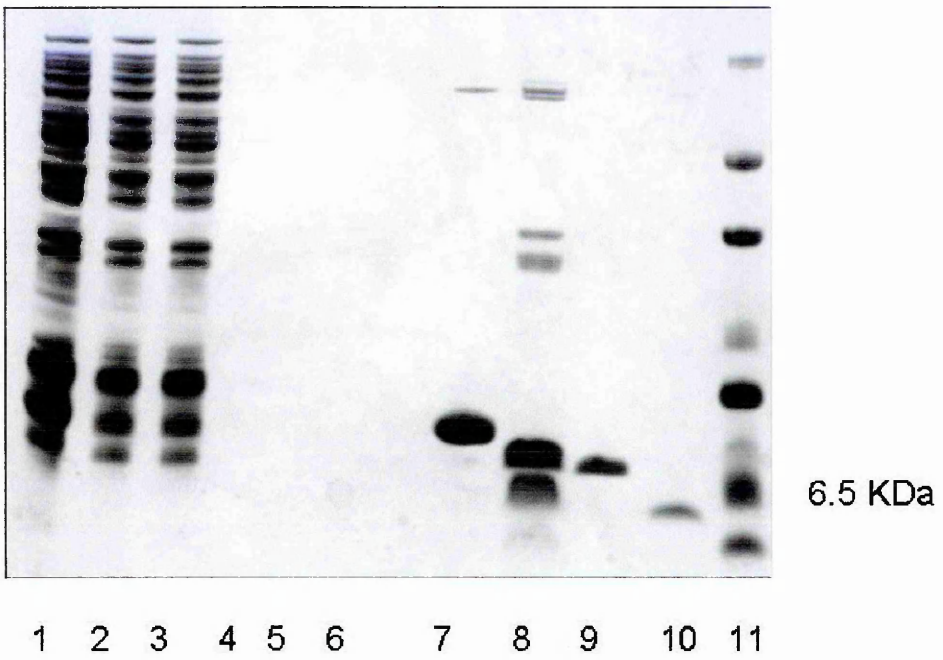


Fig. 8.2.1 10% acrylamide gel of the F1 C-terminal domain purification fractions. The lanes correspond as follows: lane1: lysed cell supernatant, lane 2: 1st load flow-through, lane 3: 2nd load flow-through, lane 4:low salt buffer 1st wash, lane 5: high salt buffer wash, lane 6: low salt buffer 2nd wash, lane 7: 300mM imidazole eluate, lane 8: Imidazole eluate after incubation with TEV, lane 9: F1 C terminus domain eluate, lane 10: 2nd 300mM imidazole wash, lane 11:molecular markers.

8.3 F1/TnH(30-61) complex production

8.3.1TnH(30-61) cloning

In order to produce ^{15}N and ^{13}C labelled TnH(30-61) in a cheaper way than custom synthesis the peptide was expressed as a fusion tail of GST. Primers spanning the full amino-acidic sequence of TnH(30-61) were designed such to have a *NcoI* and *NotI* sites at the 5' and 3' prime end respectively. An equimolar amount of the two primers was heated above the theoretical melting point in a water bath and then left to cool down slowly. The primers were then cloned into a *NcoI*/*NotI* sites of a pET24d modified vector containing GST and a TEV-protease cleavage site. The sense primers sequence is: 5´-CATGGGAAAAAAAGCGAAAAAAGGCTTTATGACCCCGGAACGCAAAAAAAA

AACTGCGCCTGCTGCTGCTGCGCAAAAAAGCGGCGGAAGAACTGAAAAAAG AATAATAAGC-3′, the anti-sense one: 5′-GGCCGCTTATTATTCTTTTTTCAGTTCTTCCGCCGCTTTTTTGCGCAGCAGCAG GCGCAGTTTTTTTTTGCGTTCCGGGGTCATAAAGCCTTTTTTTGCCGCTTTTTT TCC-3′.

8.3.2 Transformation of competent cells

E. coli strain BL21(DE3)pJY2 (Stratagene) cells were transformed with GST-TEV-TnH(30-61) plasmid DNA (1μl at 50-100 ng/ μl) using the same procedure as for F1

8.3.3 GST-TnH(30-61) fermentation

As initial test recombinant unlabelled GST-TnH(30-61) was grown in LB broth containing 30 μg/ml of kanamycin. A single colony from the LB agar plate was used to inoculate 100 ml per litre of media. The inoculation was left to grow at 37 °C until the optical density measured at 600nm (OD_{600}) was around one. The 100 ml growth was then used to inoculate one litre of media and the cell grown at 37 °C. When OD_{600} was ~0.7-0.8 isopropyl β-D-thiogalactopyranoside (IPTG) at 0.5mM final concentration was added to the cell to induce the protein production. After induction the cells are left to grow for ~3-4 h.

After the initial test isotopic labelled GST-TnH(30-61) was produced growing the cells in minimum medium supplemented with [$^{13}C_6$] glucose and/or $^{15}NH_4Cl$ as sole source of carbon and nitrogen respectively.

8.3.4 GST F1/ TnH(30-61) complex purification

In order to avoid the peptide degradation and to ease the purification procedure the peptide was co-purified with F1. First the cells where centrifuged and the pellet re-suspended in lysis buffer containing 150mM NaCl, 50mM phosphate, lysozyme, DNAse I and a tablet of 'Complete Inhibitor Cocktail' (Roche). Then not tagged pure F1 was added to the cell lysate in a roughly 1:1 ratio and the mixture left on a shaker for ~1h at 4°C. The mixture was then purified by affinity chromatography through a GST column. A column was packed with GST (1ml per litre of growth of medium) and equilibrated with

five column volumes of lysis buffer. The mixture containing the cell lysate and the added F1 was run through the column twice, then the column was washed with five column volumes of a solution containing 300mM NaCl and 50mM phosphate. These washes are designed to remove any non specific bound contaminants from the column. After the washes the F1/TnH(30-61) complex was cleaved on column from GST by treatment with TEV protease, 2-3 μl of TEV at 15 μM were used per litre of media and the mixture was gently rotated at 4°C over night. After the overnight TEV treatment the complex was eluted from the column with a 100mM NaCl, 20mM Tris buffer. The excess of TEV was removed using a column packed with Ni-NTA (0.5ml per litre of medium) The nature and purity of the proteins were evaluated by 10% acrylamide gel electrophoresis and confirmed by mass spectrometry.

8.3.5 F1/TnH/TnT complexes formation

TnH(1-224),TnH(1-340) and TnT(110-266) protein samples were kindly provided by Belinda Bullard. The samples were in 6M urea solution, the protein concentration was generally around 1.5mM.

F1 and TnH(1-224) were both dissolved in 6M urea, the complex was refolded dialyzing out urea gradually. The final ratio F1/TnH(1-224) was approximately 1/1.5 and the buffer composition was 20 mM TRIS, 200mM KCl at pH 6.8. Exactly the same procedure was followed for the refolding of all the other complexes. All the samples show to some extent precipitation and a residual amount of urea quantified in roughly 20mM. The estimate of urea content is not accurate because is done just comparing the TRIS and urea peaks in the 1D proton NMR spectra.

8.4 NMR data acquisition and processing

NMR spectra of F1 were acquired on ^{15}N or $^{5}N/^{13}C$ uniformly labelled samples at a 0.9 mM concentration in 90%/10% H_2O/D_2O and 10 mM Tris buffer (pH 6.8), 100 mM NaCl. All of the spectra were recorded at 25°C on Varian or Bruker spectrometers operating at 500, 600, and 800 MHz proton frequencies and equipped with 5 mm triple-resonance probes or cryoprobes. ^{15}N-HSQC, HNCA, HN(CO)CA, HNCO,HNCACB experiments were employed to obtain sequence specific ^{1}HN, ^{15}N, $^{13}C\alpha$, $^{13}C\beta$ and $^{13}C'$

backbone assignments (Bodenhausen and Ruben, 1980, Kay *et al.,* 1990, Bax and Ikura, 1991, Wittekind and Muller, 1993). Side chain aliphatic proton and carbon assignments were achieved by a combination of 3D ^{15}N-edited TOCSY and ^{15}N and ^{13}C edited NOESY-HSQC, H(CCO)NH and HCCH-TOCSY. (Zuiderweg and Fesik, 1989, Muhandiram *et al.,* 1993, Grzesiek *et al.,* 1993, Bax *et al.,* 1990) For the assignment of the large number (12) of aromatic side chains present in F1, (Hβ)Cβ(CγCδ)Hδ and (Hβ)Cβ (CγCδ)Hε experiments were used in combination with ^{13}C-HSQC, and HCCH-TOCSY tuned for the aromatic resonances. (Yamazaki *et al.* 1993) The WATERGATE sequence (Piotto *et al.,* 1992) was used for water suppression. The echo anti-echo and STATES-TPPI methods were used to obtain quadrature detection.(Palmer *et al.,* 1991, Kay *et al.,* 1992, Bax *et al.,* 1991) The spectra were typically processed by applying a Gaussian window function and were zero filled to double the size of the data. All the spectra were referred setting the ^{1}H water signal at 4.76 ppm at 25 °C. The spectra were processed using NMRPipe/NMRDraw (Delaglio *et al.,* 1995) and analyzed using XEASY (Bartels et al., 1995).The mixing times used for the TOCSY and NOESY type experiment were typically 60 ms and 100ms respectively.

8.5 Relaxation data

All the heteronuclear Overhauser effect (NOE), longitudinal (T_1) and transversal (T_2) relaxation rates were measured at 25 °C on a Varian Inova spectrometer operating at 500 Mhz.(Kay *et al.,* 1989, Boyd *et al.,* 1990) For the T_1 measurement twelve spectra were recorded with relaxation delay T set to: 9, 97, 194, 298, 394, 498, 595, 699, 795, 900, 996, 1196 ms. For the T_2 measurement fifteen spectra were recorded with relaxation delay T set to: 9, 17, 26, 32, 43, 52, 69, 78, 86, 95, 104, 112, 130, 138, 147 ms.

T_1 and T_2 values were obtained by non linear least squares fitting of the peaks intensities to an exponential function. Due to overlap, lack of assignment or failing in finding a good fitting only 105,79 and 67 peaks out of 158 were considered for the analysis of F1, F1/TnH(126-159) and F1/TnH(30-61) relaxation features respectively.

8.6 Structure calculation

The NOESY cross peaks for both, holo F1 and apo F1 C terminal domain, were manually peak picked and the volumes determined by the maximum integration method in Xeasy (Bartels et al., 1995). The frequency tolerance was +/-0.03 in the acquisition dimension, +/-0.04 in the indirect proton dimension and +/-0.5 in the carbon dimension for the ^{13}C edited NOESY, while it was +/-0.05, +/-0.05 and +/-0.5 respectively for the ^{15}N edited NOESY. A list of dihedral angle restrains was obtained from Talos (Cornilescu *et al* 1999).

Structure calculations were carried out with the Aria1.2 package (Linge *et al.*2003), based a standard simulated annealing CNS protocol (Brunger et al., 1998). A typical run consisted of nine iterations. The initial structure ensemble was generated at iteration 0 choosing an NOE violation tolerance and a partial assignment cut-off probability of 1000 Å and 1.01 respectively. No NOE distance restraints are excluded in this way and the partial NOE assignment is based on chemical shifts only. In the following eight iterations, as the structures improve, the NOE violation tolerances were progressively reduced (1000.0, 1.0, 0.5, 0.1, 1.0, 0.1, 0.1 and 0.1 Å), with the exception of iteration 5 in which the violation tolerance was increased to 1 Å to ensure that important NOEs consistent with the structure at that stage were not excluded. The partial assignment cut-off probability was reduced in parallel (0.9999, 0.999, 0.99, 0.98, 0.96, 0.93, 0.90 and 0.80) to eliminate ambiguous NOEs which do not contribute significantly. Twenty structures were calculated at each iteration. Floating assignment of prochiral groups and correction for spin diffusion were applied (Folmer et al., 1997; Linge et al. 2004). The best seven structures in terms of lowest global energy were selected at the end of each iteration and used in the following iteration for the assignment of additional NOEs. In the 8^{th} iteration, the structures were refined by molecular dynamics simulation in water (Linge et al. 2003). The best 10 structures in energetic terms were selected as representative of the holo F1 and apo F1 C terminal domain structures and used for statistical analysis.

8.7 NMR titration

Both peptides are highly soluble in aqueous solution, so stock solutions of both peptides were prepared. They were dissolved in 100mM KCl and 20mM TRIS and the pH was

adjusted to 6.8 using a solution 0.01N NaOH. The peptide concentrations were determined by amino acid analysis.

A solution of 0.3mM ^{15}N labelled F1 in H_2O/D_2O (9:1 v/v) was titrated with the peptide stock solutions. Both 1D and 2D ^{1}H-^{15}N HSQC NMR spectra were acquired at every titration point; the pH was checked and remained at 6.8 throughout the titration.The final ratios of F1/TnI(30-61) and F1/TnI(126-159) were 3 and 2.5 respectively.

The spectra were acquired on Varian Inova spectrometer operating at 600Mhz.

8.8 Fluorescence titration

Fluorescence emission spectra were recorded using a SPEX FluoroMax fluorimeter with excitation and emission bandwidths of 1.7 and 5 nm respectively. All measurements were made at 25°C or 40 °C in 25 mM Tris-HCl, 100 mM NaCl (pH 6.8) with added calcium (250 μM) or EDTA (1 mM) where necessary.

8.9 CD titration

Far-UV CD spectra were recorded on a Jasco J-715 spectropolarimeter. Thermal denaturation was followed by monitoring the far-UV CD signal at 222 nm whilst changing the temperature from 2°C to 95°C at 1 degree/min. Chemical denaturation was followed by monitoring the far-UV CD signal at 222 nm. Titrations were performed by adding aliquots of protein in concentrated urea (10 M) to a solution of the protein at the same concentration in aqueous buffer. All measurements were made at 20°C (unless otherwise noted) in 25 mM Tris-HCl, 100 mM NaCl (pH 6.8) with added calcium (250 μM) or EDTA (1 mM) where necessary.

8.10 UV titration

Absorbance measurements were performed on a JASCO V-550 UV/Vis spectrophotometer. All measurements were made at 25°C or 40 °C in 25 mM Tris-HCl, 100 mM NaCl (pH 6.8).

8.11 Complex structure calculation

Haddock is a protein-protein docking program that uses biochemical and/or biophysical interaction data such as, for example, chemical shift perturbation data obtained from NMR titration experiments or mutagenesis data as ambiguous interaction restraints to drive the docking.(Dominguez *et al.*, 2003) The use of these ambiguous experimental constrains narrows down the search through the conformational space of the complex geometry making the calculation less heavy, and increases the chances for the calculation to result in an unique solution. After calculation, the structures are ranked according to their intermolecular energy, that is the sum of electrostatic, van der Waals, and experimental restrains energy terms. The docking protocol consists of three stages: (i) randomization of orientations and rigid body energy minimization (EM), (ii) semirigid simulated annealing in torsion angle space (TAD-SA), and (iii) final refinement in Cartesian space with explicit solvent.

In the randomization stage, the two partner proteins are positioned at 150 Å from each other in space and each protein is randomly rotated around its centre of mass. Rigid body EM is then performed: first, four cycles of orientational optimization are performed in which each protein in turn is allowed to rotate to minimize the intermolecular energy function. Then both translations and rotations are allowed, and the two proteins are docked by rigid body EM. Typically 1000 complex conformations are calculated at this stage. The best 200 solutions in terms of intermolecular energies are then refined. The second stage consists of three simulated annealing refinements. In the first simulated annealing, the two proteins are considered as rigid bodies and their respective orientation is optimized. In the second simulated annealing, the side chains at the interface are allowed to move. In the third simulated annealing, both side chains and backbone at the interface are allowed to move to allow for some conformational rearrangements. The resulting structures are then subjected to 200 steps of steepest descent EM. The final stage consists of a refinement in water. The final structures are clustered using the pairwise backbone RMSD at the interface. A cluster is defined as an ensemble of at least two conformations displaying an iRMSD (backbone RMSD at the interface) smaller than 1.0 Å. The resulting clusters are analyzed and ranked according to their average interaction energies and their average buried surface area.

REFERENCES

Agianian B, Krzic U, Qiu F, Linke WA, Leonard K, Bullard B (2004) A troponin switch that regulates muscle contraction by stretch instead of calcium. EMBO J. 23,772-779.

Bartels, C., Xia, T., Billiter, N., Güntert, P. and Wüthrich, K. (1995) The program XEASY for computer-supported NMR spectral analysis of biological macromolecules. J Biomol. NMR 6, 1-10.

Bax A. and Ikura M. (1991) An efficient 3D NMR technique for correlating the proton and 15N backbone amide resonances with the alpha-carbon of the preceding residue in uniformly 15N/13C enriched proteins. J Biomol NMR,1(1):99-104.

Bax A., Clore G.M., Gronenborn A.M. (1990) H-1-H-1 Correlation via isotropic mixing of C-13 magnetization, a new 3-dimensional approach for assigning H-1 and C-13 spectra of C-13 enriched proteins. J Magn Reson, 88(2):425-431.

Bax A., Ikura M., Kay L.E., Zhu G. (1991) Removal of F1-base-line distortion and optimization of folding in multidimensional NMR-spectra. J Magn Reson, 91(1):174-178.

Biekofsky, R.R., Martin, S.R., Browne, J.P., Bayley, P.M., Feeney, J. (1998) Ca2+ coordination to backbone carbonyl oxygen atoms in calmodulin and other EF-hand proteins. 15N chemical shifts as probes for monitoring individual site Ca2+ coordination. Biochemistry 37, 7617-7629.

Blumenschein, T.M.A., Stone, D.B., Fletterick, R.J, Mendelson, RA. and Sykes, B.D. (2005) Calcium-dependent changes in the flexibility of the regulatory domain of troponin C in the troponin complex. J. Biol. Chem. 280,21924-21932.

Bodenhausen, G. and Ruben, D.J. (1980) Chem. Phys. Lett. 69, 185-188.

Boyd J., Hommel U., Campbell I.D. (1990) Influence of cross-correlation between dipolar and anisotropic chemical-shift relaxation mechanism upon longitudinal relaxation rates of N-15 in macromolecules. Chem Phys Lett, 175(5):477-482.

Brünger, A.T, Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., Read, R.J., Rice, L.M., Simonson, T. and Warren, G.L. (1998) Cystallography & NMR system: A new software suite for macromolecular structure determination. Acta Cryst. D 54,905-921.

Bullard B., Burkart C., Labeit S., Leonard K. (2005) The function of elastic proteins in oscillatory contraction of insect flight muscle. J.Mus Res Cell Motil 26(6-8):466-476.

Bullard B., Kevin L., Larkins A., Butcher G., Karlik C., Fyberg E. (1988) Troponin of asynchronous flight muscle. J.Mol Biol Dec5;204(3),621-637.

Cantor C.R and Schimmel P.R., *Biophysical Chemistry Part II: Technique for the study of biological structure and function.*(1980) Freeman W.H. and Company, New York.

Chandra, M., da Silva, E.F., Sorenson, MM., Ferro, J.A., Pearlstone, J.R., Nash, B.E., Borgford, T., Kay, C.M., Smillie, L.B. (1994) The effects of N helix deletion and mutant F29W on the Ca2+ binding and functional properties of chicken skeletal muscle troponin. J. Biol. Chem. 269, 14988-14994.

Cornilescu G, Delaglio F, Bax A. (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology. J. Biomol. NMR 13,289-302.

Comte M., Maulet Y., Cox J.A. (1983) Ca^{2+} dependent high-affinity complex formation between calmodulin and melittin. Biochem J. 1; 209(1): 269–272.

Creighton T.E. , *Proteins* (2), (1993) Freeman W.H. and Company, New York.

Cullen M.J. and Landon D.N. (1994) The normal ultrastruture of skeletal muscle. Disorder of Voluntary Muscle:87-134.

Davis J.S., and Rodgers M.E. (1995) Indirect coupling of phosphate release to *de novo* tension generation during muscle contraction. Proc. Natl. Acad. Sci. USA. 92:10482-10486.

Delaglio F., Grzesiek S., Vuister G.W., Zhu G., Pfeifer J., Bax A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J.Biomol NMR Nov;6(3).277-93.

Dickinson M. (2006) Insect flight. Curr Biol. 2006 May 9;16(9):309-314.

Dobbie, I., M. Linari, G. Piazzesi, M. Reconditi, N. Koubassova, M.A. Ferenczi, V. Lombardi, and M. Irving. 1998. Elastic bending and active tilting of myosin heads during muscle contraction. Nature. 396:383–387.

Dominguez C., Boelens R., Bonvin A.M. (2003) HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. J Am Chem Soc. 2003 Feb 19;125(7),1731-7.

Dvoretsky, A, Abusamhadneh, E.M., Howarth, J.W., Rosevear, P.R., (2002) Solution structure of calcium-saturated cardiac troponin C bound to cardiac troponin I. J. Biol. Chem. 277,38565-38570.

Folmer, R.H., Hilbers, C.W., Konings, R.N. and Nilges, M. (1997) Floating stereospecific assignment revisited: application to an 18 kDa protein and comparison with J-coupling data. J. Biomol. NMR 9,245-258.

Frank O. Zur Dynamik des Herzmuskels.(1895) J Biol.32:370-447.

Fredricksen, R.S. and Swenson, C.A. (1996) Relationship between stability and function for isolated domains of troponin C. Biochemistry 35, 14012-14026.

Gagné, S.M., Li, M.X., McKay, R.T. and Sykes, B.D. (1998) The NMR angle on troponin C. Biochem. Cell Biol. 76, 302-312.

Gergeley J. (1994) Biochemical aspects of muscular structure and function. Disorder of Voluntary Muscle:179-230

Gifford, J.L., Walsh M.P., Vogel H.J. (2007) Structures and metal-ion-binding properties of the Ca2+-binding helix-loop-helix EF-hand motifs. Biochem J. 2007 Jul 15;405(2):199-221.

Gordon A.M., Homsher E., Regnier M. (2000) Regulation of contraction in striated muscle. Physiol Rev. Apr;80(2),853-924.

Gordon S., Dickinson M.H., (2006) Role of calcium in the regulation of mechanical power in insect flight. Proc Natl Acad Sci USA Mar 14;103(11):120-138.

Greaser M.L., Gergely J. (1971) Reconstitution of troponin activity from three protein components. J.Biol Chem Jul 10;246(13),4226-4233.

Grzesiek S., Anglister J., Bax A. (1993) Correlation of backbone amide and aliphatic side-chain resonances in C-13/N-15- enriched proteins by isotropic mixing of C-13 magnetization. J.Magn Reson 101(1):114-119.

Gulati J., Sonnenblick E., Babu A. (1991) The role of troponin C in the length dependence of Ca(2+)-sensitive force of mammalian skeletal and cardiac muscles. J Physiol. Sep;441:305-324.

Herzberg and James (1985) Structure of the calcium regulatory muscle protein troponin C at 28 A resolution. Nature 313, 653-659.

Holmes K.C. (1995) The actomyosin interaction and its control by tropomyosin. Biophys J. 1995 Apr;68(3):2-7.

Holmes K.C., Popp D., Gebhard W., Kabsch W. (1990) Atomic model of the actin filament. Nature. 1990 Sep 6;347(6288):44-9.

Houdusse A, Love ML, Dominguez R, Grabarek Z, Cohen C (1997) Structures of four Ca^{2+}-bound troponin C at 2.0 A resolution: further insights into the Ca^{2+}-switch in the calmodulin superfamily. Structure 5, 1695-1711.

Huxley H.E. (1969) The mechanism of muscular contraction. Science Jun 20;164(886):1356-65.

Improta S., Politou A.S., Pastore A. (1996) Immunoglobulin-like modules from titin I-band: extensible components of muscle elasticity. Structure. 1996 Mar 15;4(3):323-37.

Kay L.E., Ikura M., Tschudin R., Bax A. (1990) 3-Dimensional triple-resonance NMR-spectroscopy of isotopically enriched proteins. J Magn Reson 89(3):496-514.

Kay L.E., Keifer P., Saarinen T. (1992) Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. J AM Chem Soc, 114(26): 10663-10665.

Kay L.E., Torchia D.A., Bax A. (1989) Backbone dynamics of proteins as studied by N-15 inverse detected heteronuclear NMR-spectroscopy – application to staphylococcal nuclease. Biochem, 28(6):2387-2391.

Keeler J. *Understanding NMR spectroscopy.*(2005) Wiley & Sons

Kordel, J., Skelton, N.J., Akke, M., Chazin, W.J. (1993) High-resolution structure of calcium-loaded calbindin D9k. J.Mol.Biol. v231 pp. 711-734.

Labeit S., Kolmener B. (1995) Titins: giant proteins in charge of muscle ultrastructure and elasticity. Science Oct 13;270(5234):293-296.

Laskowski, R.A., Rullman, J.A., MacArthur, M.W., Kaptein, R. and Thornton, J.M. (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. J. Biomol. NMR 8, 477-486.

Levitt M. *Spin dynamics* (2001) Wiley & Sons

Li Q., Jin J.P., Granzier H.L. (1995) The effect of genetically expressed cardiac titin fragments on in vitro actin motility. Biophys J. Oct; 69: 1508-1518.

Li, M.X., Spyracopoulos, L, Sykes, B.D. (1999) Binding of cardiac troponin-I147-163 induces a structural opening in human cardiac troponin-C. Biochemistry 38, 8289-8298.

Li M.X., Wang X., Lindhout D.A., Buscemi N., Van Eyk J.E., Sykes B.D. (2003) Phosphorylation and mutation of human cardiac troponin I deferentially destabilize the interaction of the functional regions of troponin I with troponin C. Biochemistry 16;42(49):14460-8.

Linge, J.P., Habeck, M., Rieping, W. and Nilges, M. (2004) Correction of spin diffusion during iterative automated NOE assignment. J. of Mag. Res. 167, 334-342.

Linge, J.P., Habeck, M., Rieping, W., Nilges, M. (2003) ARIA: automated NOE assignment and NMR structure calculation. Bioinformatics 19, 315-6.

Lombardi, V., G. Piazzesi, M.A. Ferenczi, H. Thirlwell, I. Dobbie, and M. Irving. 1995. Elastic distortion of myosin heads and repriming of the working stroke in muscle. Nature. 374:553–555.

Luther P.K., Squire J.M. (1980) Three-dimensional structure of the vertebrate muscle A-band. II. The myosin filament superlattice. J Mol Biol. 1980 Aug 25;141(4):409-39.

Mak AS, Smillie LB (1981) Structural interpretation of the two-site binding of troponin on the muscle thin filament. J Mol Biol 149. 541-550.

Masino, L., Martin, SR., Bayley, P.M. (2000) Ligand binding and thermodynamic stability of a multiprotein protein, calmodulin. Prot. Sci. 9, 1519-1529.

Mateos J.,Herranz R., Domingo A., Sparrow J., MarcoR. (2006) The structural role of high molecular weight tropomyosins in dipteran indirect flight muscle and the effect of phosphorylation. J Muscle Res Cell Motil 27(3-4):189-201.

Mercier P., Li, MK., Sykes, BD. (2000) Role of the structural domain of troponin C in muscle regulation: NMR studies of Ca2+ binding and subsequent interactions with regions 1-40 and 96-115 of troponin I. Biochemistry 39, 2902-2911.

Mercier P., Ferguson, R.E., Irving, M., Corrie, J.E., Trentham, D.R., Sykes, B.D. (2003) NMR structure of a bifunctional rhodamine labeled N-domain of troponin C complexed with the regulatory "switch" peptide from troponin I: implications for in situ fluorescence studies in muscle fibers. Biochemistry 42, 4333-4348.

Muhandiram D.R., Farrow N.A., Xu G.Y., Smallcombe S.H., Kay L.E. (1993) A gradient C-13 NOESY-HSQC experiment for recording NOESY spectra of C-13-labelled proteins dissolved in H2O.J Magn Reson 102(3):317-321.

Norwood T.J., Boyd J., Campbell I.D. (1989) Improved resolution in 1H-detected 1H-15N correlation experiments. FEBS Lett Sep25;255(2):369-371.

Ohtsuki L., Maruyama K., Ebashi S. (1986) Regulatory and cytoskeletal proteins of vertebrate skeletal muscle. Adv Protein Chem 38:1-67.

Ojima, T., Koizumi, N., Ueyama, K., Inoue, A., Nishita, K. (2000) Functional role of $Ca^{(2+)}$-binding site IV of scallop troponin C. J Biochem (Tokyo) 128, 803-809.

Pace C.N., Vajdos F., Fee L., Grimsley G., Gray T. (1995) How to measure and predict the molar absorption coefficient of a protein. Protein Sci. Nov;4(11):2411-23.

Palmer A.G., Cavanagh J., Wright P.E., Rance M. (1991) Sensitivity improvement in proton-detected 2-dimensional heteronuclear correlation NMR-spectroscopy. J Magn Reson, 93(1):151-170.

Peckham, M., Cripps, R., White, D. and Bullard, B. (1992). Mechanics and protein content of insect flight muscle. J. Exp. Biol. 168, 57-76.

Phan, J., Zdanov, A., Evdokimov, A. G., Tropea, J. E., Peters, H. P. K., Kapust, R. B., Li, M., Wlodawer, A., and Waugh, D. S. (2002). Structural basis for the substrate specificity of tobacco etch virus protease. J. Biol. Chem. 277: 50564-50572.

Philips G.N., Chacko S. (1996) Mechanical properties of tropomyosin and implications for muscle regulation. Biopolymers Jan;38(1):89-95.

Piazzesi, G., M. Linari, M. Reconditi, F. Vanzi, and V. Lombardi. 1997. Cross-bridge detachment and attachment following a step stretch imposed on active single frog muscle fibres. J. Physiol. 498:3-15.

Piotto M., Saudek V., Sklenar V. (1992) Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. J.Biomol NMR Nov;2(6):661-5.

Potter, J.D., Sheng, Z., Pan, B.S., Zhao, J. (1995) A direct regulatory role for troponin T and a dual role for troponin C in the Ca^{2+} regulation of muscle contraction. J. Biol. Chem. 270, 2557-2562.

Pringle JW (1949) The excitation and contraction of the flight muscle of insects.J Physiol 108, 226-232.

Pringle JW (1978) The Croonian Lecture, 1977. Stretch activation of muscle: function and mechanism. Proc R Soc Lond B Biol Sci. 1978 May 5;201(1143):107-130.

Qiu.F., Lakey A, Agianian B, Hutchings A, Butcher GW, Labeit S, Leonard K, Bullard B. (2003) Troponin C in different insect muscle types: identification of two isoforms in *Lethocerus*, *Drosophila* and *Anopheles* that are specific to asynchronous flight muscle in the adult insect. Biochem J. 371, 811-821.

Satyshur, K.A., Rao, S.T., Pyzalska, D., Drendel, W., Greaser, M., Sundaralingam, M. (1988) J. Biol Chem. 263, 1628-1647.

Sia, S.K., Li, M.X., Spyracopoulos, L., Gagné, S.M., Liu, W., Putkey, J.A., Sykes, B.D. (1997) Structure of cardiac muscle troponin C unexpectedly reveals a closed regulatory domain. J. Biol. Chem. 272, 18216-18221.

Slupsky, C.M. and Sykes, B.D. (1995) NMR solution structure of calcium saturated skeletal muscle troponin C. Biochemistry 34, 15953-15964.

Spyracopoulos L., Gagne L.M., Li M.X., Sykes B.D. (1998) Dynamics and thermodynamics of the regulatory domain of human cardiac troponin C in the apo- and calcium-saturated states. Biochemistry. 1998 Dec 22;37(51):18032-18044.

Squire J.M.(1981) The structural basis of Muscular Contraction. New York: Plenum Press

Starling EH. Linacre Lecture on the Law of the Heart.(1918) London, England Longmans.

Starling EH and Visscher MB. The regulation of the energy output of the heart.(1926) J Physiol 62: 243–261.

Steiger G.J. (1971) Stretch activation and myogenic oscillation of isolated contractile structures of heart muscle. Pflugers Arch. 330:347-361.

Steiger, G.J. (1977) Tension transients in extracted rabbit heart muscle preparations. J. Mol. Cell. Cardiol. 9:671–685.

Skelton, N.J., Kordel, J., Chazin, W.J. (1995) Determination of the solution structure of Apo calbindin D9k by NMR spectroscopy. J.Mol.Biol. v249 pp. 441-462.

Takeda S., Yamashita A., Maeda K., Maeda Y. (2003) Structure of the core domain of human cardiac troponin in the Ca(2+)-saturated form. Nature. 2003 Jul 3;424,35-41.

Tobacman L.S. (1996) Thin filament-mediated regulation of cardiac contraction. Annu Rev Physiol. 1996;58:447-81.

Tripet, B., van Eyk, J.E., Hodges, R.S. (1997) Mapping of a second actin-tropomyosin and a second troponin C binding site within the C terminus of troponin I, and their importance in the Ca^{2+}-dependent regulation of muscle contraction. J. Mol. Biol. 271, 728-750.

Tripet, B., de Crescenzo, G., Grothe, S., O'Connor-McCourt, M., Hodges, R.S. (2003) Kinetic analysis of the interactions between Troponin C (TnC) and Troponin I (TnI) binding peptides: evidence for separate binding sites for the 'structural' N-terminus and the 'regulatory' C-terminus of TnI on TnC. J. of Mol. Recognit. 16, 37-53.

Trinik J. Cytoskeleton titin as a scaffold and a spring. (1996) Curr Biol 6:258-289.

Tsuda, S., Miura, A., Gagné, S.M., Spyracopoulos, L., Sykes, B.D. (1999) Low-temperature-induced structural changes in the apo regulatory domain of skeletal muscle troponin C. Biochemistry 38, 5693-5700.

Vassylyev, D.G., Takeda, S., Wakatsuki, S, Maeda, K., Maeda, Y. (1998) Crystal structure of troponin C in complex with troponin I fragment at 2.3-A resolution. Proc. Natl. Acad. Sci. USA 95, 4847-4852.

Vemuri R., Lankford E.B., Poetter K., Hassanzadeh S., Takeda K., Yu Z.-X, Ferrans, V.J., Epstein N.D. (1999) The stretch-activation response may be critical to the proper functioning of the mammalian heart. Proc. Natl. Acad. Sci. USA. 96:1048–1053.

Vibert, P. Craig, R., Lehman, W. (1997) Steric-model for activation of muscle thin filaments. J. Mol. Biol. 266, 8-14.

Vinogradova M.V., Stone D.B., Malanina G.G., Karatzaferi C., Cooke R., Mendelson R.A., Fletterick R.J. (2005) Ca(2+)-regulated structural changes in troponin. Proc Natl Acad Sci U S A. Apr 5;102(14),5038-5043.

Vigoreaux J.O.(1994) The muscle Z-band: lesson in stress management. J.Mus Res Cell Motil 15:237-255.

Vriend, G. (1990) WHAT IF: a molecular modelling and drug design program. J. Mol. Graph. 8, 52-56.

Ward D.G., Brewer S.W., Gallon C.E., Gao Y., Levine B.A., Trayer I.P.(2004) NMR and mutagenesis studies on the phosphorylation region of human cardiac troponin I. Biochemistry. 2004 May 18;43(19):5772-81.

Wittekind M. and Muller L. (1993) HNCACB, a high sensitivity 3D experiment to correlate amide –proton and nitrogen resonances with the alpha-carbon and beta-carbon resonances in proteins. J Magn Reson, 101(2): 201-205.

Wray J.S. (1979) Filament geometry and the activation of insect flight muscles. Nature 280: 325-326.

Yamazaki T., Forman-Kay J.D., Kay L.E. (1993) Two-dimensional NMR experiments for correlating $^{13}C\beta$ and $^{1}H\delta/\varepsilon$ chemical shifts of aromatic residues in ^{13}C-labeled proteins via scalar couplings. J Am Chem Soc 115:11054–11055.

Yap, K.L., Ames, J.B., Swindells, M.B., Ikura, M. (1999) Diversity of conformational states and changes within the EF hand protein superfamily. Proteins 37, 499-507.

Zhang M., Tanaka T., Ikura M. (1995) Calcium-induced conformational transition revealed by the solution structure of apo calmodulin. Nat Struct Biol.;2(9):758-67.

Zuiderweg E.R.P. and Fesik S.W. (1989) Heteronuclear 3-dimensional NMR spectroscopy of the inflammatory protein C5A. Biochem, 28(6):2387-2391.

Zwalen C., Legault P., Vincent J.F.S., Greenblatt J., Konrat R., Kay L.E. (1997) Methods for measurement of intermolecular NOEs by multinuclear NMR spectroscopy: application to a bacteriophage λ N-peptide/*box*B RNA complex. J Am Chem Soc 119:6711-6721.